AFOSR-TR- 80 - 0336

ADA084739

THE COLLEGE OF EARTH AND MINERAL SCIENCES

APPLICATION OF TWO MULTIVARIATE CLASSIFICATION
TECHNIQUES TO THE PROBLEM OF
SEISMIC DISCRIMINATION

A Paper in Geophysics

by

Alan G. R. Bell

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 1978

THE PENNSYLVANIA STATE UNIVERSITY

UNIVERSITY PARK, PENNSYLVANIA

THE PENNSYLVANIA STATE UNIVERSITY

COLLEGE OF EARTH AND MINERAL SCIENCES

UNDERGRADUATE PROGRAMS OF STUDY:
Ceramic Science, Earth Sciences, Geography, Geological Sciences, Metallurgy, Meteorology, Mineral Economics, Mining, Petroleum and Natural Gas, Polymer Science

GRADUATE PROGRAMS AND FIELDS OF RESEARCH:
Ceramic Science, Fuel Science, Geochemistry, Geography, Geology, Geophysics, Metallurgy, Meteorology, Mineral Economics, Mineral Processing, Mineralogy and Petrology, Mining Engineering, Petroleum and Natural Gas Engineering

INTERDISCIPLINARY GRADUATE PROGRAMS:
Earth Sciences, Environmental Pollution Control, Mineral Engineering Management, Solid State Science

ASSOCIATE DEGREE PROGRAMS:
Materials Technology
Mining Technology

INTERDISCIPLINARY RESEARCH SECTIONS:
Coal Research, Mass Spectrometry, Mine Drainage, Mineral Conservation, Ore Deposits

ANALYTICAL AND STRUCTURE STUDIES:
Classical Chemical Analysis of metals and silicate and carbonate rocks, X-Ray Crystallography, Electron Microscopy and Diffraction, Electron Microprobe Analysis, Atomic Absorption Analysis, Spectrochemical Analysis

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) Unclassified

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR-80-0336	2. GOVT ACCESSION NO.: AD-A084 739
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): APPLICATION OF TWO MUTIVARIATE CLASSIFICATION TECHNIQUES TO THE PROBLEM OF SEISMIC DISCRIMINATION	5. TYPE OF REPORT & PERIOD COVERED: INTERIM
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Alan G R Bell	8. CONTRACT OR GRANT NUMBER(s): AFOSR-78-3528, ARPA Order-3291
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Geosciences, Pennsylvania State University, University Park, PA 16802	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: A.0.3291 8F10 61101E
11. CONTROLLING OFFICE NAME AND ADDRESS: ARPA/NMR, 1400 Wilson Blvd., Arlington, VA 22209	12. REPORT DATE: 11 Aug 78
	13. NUMBER OF PAGES: 175
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): AFOSR/NP, Bolling AFB, Bldg. #410, Wash DC 20332	15. SECURITY CLASS. (of this report): unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

9 Master's thesis

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Two multivariate classification methods have been applied to discriminant variables measured on a set of 20 Nevada Test Site (NTS) underground nuclear explosions and 27 North American Earthquakes. The classification methods used in this study are cluster analysis, and linear and quadratic multiple discriminant analysis. Discriminant analysis requires that the parameters of the underlying stochastic model be specified a priori, which is usually accomplished with a sample training set of events. Cluster analysis, however,

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

408103

follows a non-parametric approach and requires no prior information. The discriminant variables are the seven energy ratios calculated for these events by Booker and Mitronovas (1964). An additional discriminant variable, M(s)/m(b) was estimated for a reduced set of nine earthquakes and 21 underground explosions. Four M(s)/m(b) estimates were obtained for each event in this subset by randomly generating M(s) estimates within four specified variance levels about the M(s) versus M(b) lines given by Alexander and Lambert (1973). The variances used in generating the M(s) estimates are multiples of the variance found by von Seggern (1972) to be representative of M(s) measurements for NTS explosions.

The Pennsylvania State University

The Graduate School

Department of Geosciences

Application of Two Multivariate Classification Techniques to the Problem of Seismic Discrimination

A Paper in

Geophysics

by

Alan G.R. Bell

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 1978

The signatories below indicate that they have read and approved the paper of Alan G.R. Bell.

Date of Signature:	Signatories:
July 24, 1978	Shelton S. Alexander S.S. Alexander, Professor of Geophysics, Chairman of Geophysics Graduate Program, Thesis Advisor
July 21, 1978	R J Greenfield R.J. Greenfield, Associate Professor of Geophysics, Graduate Faculty Reader
June 15, 1978	P.M. Lavin, Associate Professor of Geophysics, Graduate Faculty Reader
June 19, 1978	C.A. Langston, Assistant Professor of Geophysics, Graduate Faculty Reader

ABSTRACT

Two multivariate classification methods have been applied to discriminant variables measured on a set of 20 Nevada Test Site (NTS) underground nuclear explosions and 27 North American Earthquakes. The classification methods used in this study are cluster analysis, and linear and quadratic multiple discriminant analysis. Discriminant analysis requires that the parameters of the underlying stochastic model be specified a priori, which is usually accomplished with a sample training set of events. Cluster analysis, however, follows a non-parametric approach and requires no prior information. The discriminant variables are the seven energy ratios calculated for these events by Booker and Mitronovas (1964). An additional discriminant variable, M_s/m_b was estimated for a reduced set of nine earthquakes and 21 underground explosions. Four M_s/m_b estimates were obtained for each event in this subset by randomly generating M_s estimates within four specified variance levels about the M_s versus m_b lines given by Alexander and Lambert (1973). The variances used in generating the M_s estimates are multiples of the variance found by von Seggern (1972) to be representative of M_s measurements for NTS explosions.

Where the two multiple discriminant analysis methods were used to classify events on the basis of the seven

energy ratios, two events were misclassified in both cases. The average posterior probability of correct classification was 0.79 for the linear analysis and 0.95 for the quadratic analysis. For the reduced event set where the M_s/m_b estimates were included, no events were misclassified in the quadratic analysis and only one event was misclassified in the linear analysis. This single misclassification occurred where the widest variance set of M_s/m_b values was used. In all cases where the M_s/m_b estimates were included, the average posterior probability of correct classification was found to be greater than or equal to 0.96.

Application of cluster analysis methods to the data failed to produce homogeneous groupings at the two-cluster level, both in the case where the seven energy ratios were used alone, and for the reduced set of events where the M_s/m_b ratios were included. However, where the M_s/m_b estimates were used, three homogeneous explosion clusters and one earthquake cluster with a single misplaced explosion were formed at the four-cluster level. The earthquake cluster then fused with an explosion cluster, so that homogeneous clusters did not persist to the two-cluster level. In the case where only the seven energy ratios were used, inhomogeneous clusters were formed at the earliest steps in the clustering process.

Suggestions have been made for improving discrimination capabilities based on recent advances in seismic instrumentation and data processing techniques.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd.)

TABLE OF CONTENTS (cont'd.)

LIST OF FIGURES

LIST OF FIGURES (cont'd.)

LIST OF FIGURES (cont'd.)

LIST OF TABLES

ACKNOWLEDGMENTS

The author would like to express his most sincere appreciation to Dr. S. S. Alexander for his guidance and many helpful suggestions which made possible the completion of this research.

Computing facilities for this research were provided by the Pennsylvania State University Computation Center.

This research was supported by Grants AFOSR-78-3528 and AFOSR-73-2515 under sponsorship of the Advanced Research Projects Agency.

CHAPTER I

INTRODUCTION

The need to limit the testing of nuclear devices became evident in the mid 1950's, sparked by a growing concern in the scientific community over the increasing amounts of radioactivity accumulating in the biosphere. Public awareness began to focus on this issue in 1954, after the United States exploded a 15 megaton device, Bravo, in an underwater test at Eniwetok Atoll in the Bikini Islands. Subsequent fallout from this test had been much greater than anticipated and resulted in the severe exposure of a number of Marshall Islanders to large doses of radiation. This sequence of events reached a climax when, two weeks after the explosion, a Japanese fishing vessel, The Lucky Dragon, docked at Yaizu Harbor with twenty three members of its crew suffering from radiation sickness (Lapp, 1958). These events, combined with a growing scientific awareness of the effect of radioactive fallout made it clear that some international agreement needed to be reached concerning the future of nuclear weapons testing.

The 1958 Geneva Conference of Experts

The first international political action to evaluate the consequences of nuclear testing and to study ways in which a realistic test ban treaty could be implemented took place at Geneva in 1958. The primary task of this conference, which included delegates from the United States and the Soviet Bloc, as well as those from Britain, France, and Canada, was to determine whether existing methods of detection were sufficient to ensure a test ban agreement. By July 4, 1958, an agenda had been adopted, indicating four areas of discussion. These were (a) an exchange of opinions on the methods for detecting nuclear explosions, (b) determination of methods that would reliably indicate that an explosion has taken place, (c) agreement on a system for controlling the observance of a test ban treaty, and (d) preparation of a report setting forth the expert conclusions and recommendations for such a control system (Bolt, 1976).

It was soon agreed that existing methods were adequate for detecting an atmospheric or surface test, because such an explosion generates an atmospheric pressure front which is easily detected on microbarograph recordings. In addition, radioactive debris and electromagnetic disturbances generated by a nuclear explosion on the surface or in the atmosphere serve as reliable indications that such

an explosion has taken place. Prompted by these results, the delegates to the conference issued the following statement (Bolt, 1976):

> "... the sensitivity of modern physical, chemical, and geophysical methods of measurement makes it possible to detect nuclear explosions at considerable distances. Thus, it is known that explosions of high yield which are set off on the surface of the earth and in the lower part of the atmosphere can be detected without difficulty at points on the globe that are very remote from the site of the explosion."

On the other hand, agreement on suitable methods for detecting underground tests appeared to be much more difficult to reach. The Soviet delegation was confident that existing detection methods were adequate to police a test ban agreement; however, Western delegates did not share this optimism. The debate centered on the determination of the minimum magnitude threshold at which an explosion can be detected. The Russian opinion held that the vast majority of recorded events could be unambiguously identified as earthquakes. For example, deep shocks of magnitude 5.0 or greater could be excluded, as could events occurring near control stations, populated areas, and at great depths under the ocean. This would leave only a very small number of questionable events, perhaps five or ten in a year, which would require further investigation.

The British position, presented by Sir Edward Bullard, questioned this optimism. Bullard asserted that the Russian arguments were based on data obtained under ideal

conditions, whereas an actual control network would be required to operate under realistic conditions of noise, and certainly without forewarning of impending tests. Bullard further suggested that magnitude/yield detection thresholds might be reduced by incorporating the polarity of first motion as a discriminant criterion. This criterion became quite attractive to the Geneva delegates because it was believed to represent a positive identifier of explosions. That is, a certain pattern of first motions, one that is compressional at all stations, is both a necessary and sufficient characteristic to indicate that the event in question is an explosion. The first motion polarity criterion led to a great deal of optimism among both Eastern and Western delegates, who felt that 90 percent of all continental earthquakes could be reliably distinguished from explosions on the basis of the first motion criterion alone.

Several other possible approaches to discrimination were also discussed. Most earthquakes are found to occur at depths appreciably greater than five to eight kilometers, generally considered to be the limit to which an emplacement hole for a nuclear device can be drilled. The Geneva delegates agreed that the focal depth could be reliably determined using the pP-P interval time. Under this plan, if an event was found to have a focal depth less than some threshold depth, say ten kilometers, it would be classified as a suspicious event. Similarly, the location process itself was suggested as a discrimination technique, where

events occurring in an aseismic region would be considered suspicious, and those occurring near population centers, or under deep oceans would be classified as natural events.

The third discrimination technique considered at Geneva is based on theoretical differences in the partitioning of energy between compressional and shear wave radiation fields. The source function for an underground explosion may be approximated by a pressure pulse applied to the surface of a spherical cavity, centered about the detonation point of the explosion. This type of impulsive, radially symmetric source generates only compressional waves in an infinite medium; however, boundary interactions occurring in a halfspace or a multilayered medium result in the generation of Rayleigh surface waves and in the conversion of some of the P*wave energy into SV*waves at the boundaries. The earthquake source, on the other hand, is a faulting mechanism which may be modelled by a series of double couples. This type of source entails release of shear strain energy which results in a significant partitioning of energy into shear wave radiation. In addition, where the propagating medium may be represented by one or more layers over a halfspace, Love surface waves may be generated because of the horizontally polarized (SH) shear wave components in the earthquake radiation field.

Based on these results, the seismic signature of an underground nuclear explosion should contain a large fraction of compressional wave energy and no SH or Love wave

energy, whereas the signature for an earthquake should contain a larger amount of both SV and SH shear wave energy, and both Love and Rayleigh surface wave energy. This fundamental difference in energy partitioning at the source forms the basis for the M_s/m_b discriminant, and for the cluster and discriminant analysis schemes presented in this paper.

The Geneva Control Scheme

After considering the various diagnostic techniques to be applied to the discrimination problem, the delegates to the Geneva conference turned their attention to the question of a suitable worldwide control system to monitor a subsequent treaty agreement. The Russian proposal consisted of a network of approximately 100 acoustic monitoring stations to detect surface and atmospheric blasts, coupled with existing national seismographic stations to be used for underground detection purposes. Such a system, according to the Soviet delegation, would have the capability of detecting underground and above-ground detonations to a one kiloton threshold yield (Bolt, 1976). Objections to this proposal were voiced by Western delegates, who pointed out that many of the national stations were untrustworthy, in that they were poorly equipped, and only spottily staffed. At that time a wide variety of equipment was in use, with

different frequency bandwidths and instrument response characteristics. In particular, very few stations had accurate means for time calibration, so that the variance in origin time as determined by such a network might be as great as several minutes.

The United States' plan for identifying explosions at the one kiloton yield threshold based on first motion polarity required a total of 650 stations world-wide. This proposal, in turn, proved untenable to the Eastern side, so that a deadlock ensued until a compromise proposal was suggested. British delegate William Penny proposed that the minimum detection yield threshold be raised from one kiloton to five kilotons, which would allow for a reduction in the required number of monitoring stations from 650 to 170. This suggestion evolved to form the basis for the Geneva Seismic Detection Network.

The so-called Geneva Network stands as a landmark for seismology; although it was never put into operation, it was to become the basis for a standardized global network. In all, 170 stations were to be located in North America, Europe and Asia, with some stations placed on islands in the Pacific Ocean and Carribbean Sea. In addition to the land-based stations, 10 acoustic monitoring stations were to be placed on ships. The standardized equipment was to consist of two short-period horizontal and ten short-period vertical instruments distributed over two or three kilometers in a small aperture array. Long-period instruments were also to

be included at each station, with hydroacoustic equipment included at the island stations (Bolt, 1976).

One of the major factors preventing the implementation of a comprehensive test ban treaty is the question of on-site inspection. At the 1958 Geneva Conference, this was a key provision for the West, that the sites of so-called "residue events", that is, events which could not be unambiguously identified as natural earthquakes, be open to inspection by all concerned parties. This proposal proved to be most unpalatable to the Eastern delegation, and continues to be a major stumbling block, even in present negotiations.

Because of this inability to reach agreement on the fundamental question of on-site inspection, and in light of some rather disturbing developments revealed in the October, 1958 Hardtack II tests, plans for a comprehensive test ban treaty were temporarily suspended, and the Geneva conference adjourned. Upon examination of the records from Hardtack II, it was found that several stations had recorded the earlier Rainier blast, the model used in the Geneva talks, with an anomalously high amplitude. This led to a downward revision of the magnitude for Rainier from 4.6 to 4.1. Because so many more earthquakes occur at this smaller magnitude each year, nearly twice as many events would have to be examined for a given magnitude yield threshold. In addition, it was found that the criterion of first motion polarity, then thought to be the only positive

identification criterion available, was reliable only at noise levels much lower than was previously thought necessary.

These discoveries brought home the need for massive improvements in seismic techniques. In the United States, new funding was provided for seismic improvement under project Vela Uniform, and similar research and development operations were established abroad, perhaps most notably the continuing research effort at Blacknest in Great Britain.

Under project Vela Uniform, initiated by the Advanced Research Projects Agency of the Department of Defense, 250 million dollars were spent to improve seismic capability in the United States and abroad between 1960 and 1971. Some of this funding was directed toward seismological research although much of it was required to upgrade detection capabilities and instrumentation. For example, it was during this period that the World Wide Standard Seismograph Network (WWSSN) was established, as well as numerous seismic arrays, such as the Large Aperture Seismic Array (LASA), the Alaskan Long Period Seismic Array (ALPA), and the Norwegian Seismic Array (NORSAR).

Further Negotiations and Treaties

Between 1958 and 1963, considerable progress had been achieved in detection of blasts in the atmosphere and

underwater. In these instances, the problem is considerably simplified by the fact that the detonation medium is an acoustic medium, without most of the complex propagation characteristics of the solid earth. Particularly important is the fact that neither the atmosphere nor the oceans have a competing background of natural events of similar character and magnitude.

The reverse is true in the case of seismological detection of underground explosions, where a perpetual background of natural seismicity is present, so that an effective method for differentiating between an underground explosion and a natural earthquake is required.

In light of these results, a limited test ban treaty was developed through negotiations at Geneva in 1963. In effect, the treaty bans all atmospheric and underwater testing, and severely resticts surface detonations by requiring that radioactive debris be contained entirely within the territorial limits of the nation in which the explosion is detonated.

Weapons testing has been further restricted in subsequent treaties banning tests in Antarctica (1959), the Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space (1967), and the Treaty on the Non-Proliferation of Nuclear Weapons (1968). However, underground testing has continued unabated since the ratification of the Limited Test Ban Treaty in 1963.

In 1974, India became the sixth nuclear power (the

other five are the United States, the Soviet Union, Great Britain, France, and the Peoples' Republic of China). This event reawakened concern over the proliferation of nuclear weapons and renewed the apparently long dormant international interest in banning, or at least limiting, the underground testing of nuclear weapons. As a direct consequence of the Indian nuclear test, the Threshold Test Ban Treaty was drafted in 1974. This treaty was initialled by both the United States and the Soviet Union, and although the treaty was never brought before the United States Senate for ratification, both sides informally agreed to abide by its provisions. In essence, the 1974 Limited Test Ban Treaty stipulates that underground nuclear testing be limited to devices of 150 kiloton yield or less, well above the threshold detection and identification capabilities of either side.

With the continued Soviet opposition to on-site inspection, most research since 1958 has focused on those detection and discrimination criteria which are applicable at teleseismic distances. However, recent talks in Geneva indicate the possibility that a comprehensive test ban treaty may be drafted in the near future. This has been brought about by the apparent willingness of each side to allow a certain number of stations to be installed and operated within its national borders by the other. Given this possibility of a network of internal monitoring stations, it has become imperative to develop regional

discrimination techniques to complement current teleseismic methods and better ensure compliance with a test ban treaty. Research continues on seismic discrimination, and regional techniques which have received relatively little attention in the past are now quickly growing in importance, especially in light of the recent advances in seismic instrumentation. A primary example is the recently developed Seismic Research Observatory (SRO) Network, where data is automatically digitized and stored on magnetic tape. Three outstanding features of this network are its wide dynamic range, excellent signal-to-noise characteristics, and on-site digital recording which allows the data to be routinely subjected to sophisticated digital analyses.

The remainder of this paper is devoted to describing various regional discriminant criteria and to illustrating several practical techniques by which these criteria may be employed. Among these techniques are multivariate linear and quadratic discriminant analysis, and several types of cluster analysis. These techniques are applied to the set of energy ratios measured on a group of earthquakes and underground nuclear explosions examined originally by Booker and Mitronovas (1964). This data set is augmented with estimated M_s/m_b ratios which have been generated in such a way as to reflect the uncertainties typical of western United States earthquakes and Nevada Test Site explosions.

The last part of this paper contains the description of

a general spectral analysis program which may be used to calculate spectral discriminants based on the event's spectrogram, a two-dimensional contour display of power spectral density as a function of both time (or velocity) and frequency. Alternatively, the program may be used to generate discriminants based on the spectra of time or velocity windows chosen at any position within the record. An interesting feature of this program is that spectra for long windows are calculated by averaging the spectra of the shorter windows used to obtain the event spectrogram, so that only one set of power spectra need be calculated. When coupled with the digital recording capabilities of the SRO Network, criteria such as the spectral and energy ratios described above may be routinely and economically applied. This approach provides a practical means of using the entire seismic signature in a quantitative manner for improved discrimination capabilities.

CHAPTER II

A BACKGROUND FOR THE DISCRIMINANTS USED IN THIS STUDY

The discriminants which might be employed in an earthquake-underground nuclear explosion discrimination scheme may be divided into two groups. The first group consists of those discriminants which indicate the origin of a seismic event on the basis of geographic considerations. For example, an event located at great depth or close to a major population center may be identified as a natural earthquake with a high degree of certainty. On the other hand, a shallow event occurring in a comparatively aseismic region may be regarded as suspicious. Where adequate station coverage is available so that accurate seismic location and depth determinations may be routinely obtained by seismological cataloging agencies, these geographic discriminants may be readily utilized to classify the majority of recorded events.

Those events which cannot be classified using simple geographic criteria must then be subjected to more sophisticated analyses, based on theoretical or empirical differences between the earthquake and underground explosion source characteristics. Factors contributing to these

differences are the rupture mechanism, the fault dimensions and geometry, and the source time history. Considerable effort has been made in recent years to model the earthquake and explosion source mechanisms on the basis of observed radiation patterns and to justify these models physically. As a result of this research, numerous discriminants have been developed to differentiate between earthquakes and underground explosions on the basis of the observable features of their respective radiation fields.

A Brief Survey of Earthquake and Underground Explosion Source Theory

One of the most complex and difficult problems in seismology is the development of a source theory to explain in detail the features of the observed seismic signatures of earthquakes and underground explosions. Several hypotheses have been advanced concerning the nature of the seismic focal mechanism, some of which are quite successful in explaining major characteristics of the observed radiation fields.

One of the earliest attempts to explain the earthquake source mechanism is the elastic rebound theory of fracture formulated by Reid (1933). Based on observations of the 1906 San Francisco earthquake, Reid concluded that an earthquake occurs as the result of the accumulation of elastic stresses

within a given region of the Earth. These stresses slowly increase with time until they exceed the strength of the medium, at which point the accumulated stresses are relieved by fracture. This theory may be summarized in the following statements (Stacey, 1969):

1. The earthquake is produced by fracture which is the result of elastic strains that exceed the strength of the material. These strains are produced by relative displacements of neighboring portions of the Earth's crust.

2. The relative displacements are not produced suddenly at the time of fracture, but accumulate slowly over a period of time.

3. The only mass movements at the time of fracture are due to the sudden elastic rebound of the sides of the fracture toward equilibrium. These displacements extend only a few kilometers in a direction normal to the fracture surface.

4. Elastic waves propagate from the sides of the fracture, the surface of which has, at first, a small area. This area may become large, but not at a rate greater than the compressional wave velocity of the material (Note: There has been some debate in recent years concerning the speed of rupture propagation. Some authors (Press, et al., 1961) believe the rupture velocity to be less than the S-wave velocity of the medium.)

5. The energy liberated in an earthquake is stored prior to fracture in the form of elastic strain energy.

By examination of the distribution of first motion polarities, it was found that identical first motion patterns could be obtained by representing the source in terms of an equivalent point source (Stauder, 1962; Hodgson and Stevens, 1964; Keilis-Borok, 1960). Honda (1961) considered both the single couple and the double couple point source representations, referring to them as the type

I and type II sources, respectively. He found that the double couple representation is equivalent to a system of forces acting within a spherical volume element where the net moment is balanced.

Both the single and double couple point source representations result in an identical pattern of P-wave first motions polarities. However, the S-wave radiation pattern resulting from a single couple source is twin-lobed, whereas that resulting from the double couple source is four-lobed. Studies of earthquake first motion radiation patterns (Stauder, 1962; Stevens, 1969) indicate that the double couple point source is generally the more accurate representation for the earthquake source. Moreover, the single couple point source entails a net moment which implies a rotation of the fault surface. This phenomenon is rarely observed for shallow events; however, Chinnery (1960) has suggested that plasticity effects might account for single couple type radiation from deeper events.

Ben-Menahem (1961) studied surface wave excitation using a finite line distribution of directed point sources, and simulated the effect of a propagating fracture in a half-space by integrating the elementary line source solutions assuming a finite rupture velocity. Harkrider (1964) and Ben-Menahem and Harkrider (1964) studied surface wave excitation in a layered earth model, using the directed point source to form couples, double couples, and higher order multipole source terms.

The rupture mechanism involved in the elastic rebound theory was originally thought to be that of simple Coulomb fracture. However, Coulomb fracture provides a plausible explanation only in the case of very shallow earthquakes. Because the coefficient of internal friction for most rock materials is on the order of unity, frictional sliding may occur only if the shearing stress is of the same order of magnitude as the normal stress (roughly equivalent to the overburden pressure). However, the overburden pressure increases on the average by about 300 bars per kilometer in depth, whereas the shear strength in the crust is thought to be on the order of a few hundred bars in magnitude, and even less in the upper mantle (Stacey, 1969). Thus at depth, accumulation of shear stress should result in bodily deformation (creep) rather than frictional sliding. This argument is nullified somewhat by the introduction of large hydrostatic pore pressures, such as those resulting from metamorphic processes, which reduce the effective normal stress across the fracture surface. It is thought, however, that even this modified Coulomb fracture is not a major contributing factor at depths greater than 25 kilometers (Stacey, 1969).

Attempts to explain the physics of the rupture phenomenon have resulted in two fairly successful earthquake models. Steketee (1958) demonstrated the applicability of elastic dislocation theory to the earthquake source problem. In this case, rupture formation is equivalent to the

instantaneous removal of stresses across the fault surface, which is also equivalent to the sudden application of stresses in an unstressed medium (Turnbull, 1976).

Equivalent point sources for the dislocation model have been considered by Chinnery (1960,1961), Knopoff and Gilbert (1960), and Burridge and Knopoff (1964). Based on equilibrium considerations, Chinnery (1960) concluded that the double couple representation is the most likely for near-suface faults. In the case of the single couple, the presence of an unbalanced moment implies that the medium readjusts to equilibrium at some time after the first emission of seismic waves. Chinnery (1960) suggested that such a situation might occur at depth if plasticity effects in the rupture zone are taken into account. Knopoff and Gilbert (1960) modelled the earthquake fault plane in terms of a geometric discontinuity in strain or displacement, and derived the double couple representation for the displacement discontinuity. They also found that a discontinuity in shear strain is equivalent to a directed point force representation, and that shear strain release in a laminar region (en echelon faulting) may be represented by the single couple source.

The point source dislocation model has been used to estimate the fault length, stress drop, and seismic moment from observed radiation fields based on theoretical results obtained by Brune (1970), and Hanks and Wyss (1972).

Archambeau (1968) has developed a rather formidable

source theory where the rupture process is modelled in terms of a volume relaxation phenomenon, described mathematically as an initial value problem. In this theory, the dynamical properties of the source are described in terms of the relaxation or readjustment of pre-existing stress fields. The initial shear stress field is taken to be the primary cause of rupture. Seismic effects are the result of the relaxation of this prestress shear field in the medium surrounding a region whose physical properties, rigidity in particular, have changed suddenly.

The tectonic source is viewed in terms of the release of potential strain energy from within a nonelastic zone in the medium. Within this nonlinear rupture zone, stress release is achieved by flow, fracture, or phase change, so that strain energy within the rupture volume is dissipated in the work of nonelastic deformation.

This dynamic relaxation model requires that the stress field surrounding the rupture zone readjust to equilibrium in a fashion determined by the boundary conditions on the rupture surface once rupture is initiated. This readjustment is accomplished by seismic radiation wherever the medium is elastic.

The radiation field may be characterized as a multipole expansion. For separable source-time functions, the multipole coefficients are dependent only upon source geometry. For nonseparable source-time functions, however, the multipole coefficients are frequency dependent

(Turnbull, 1976).

A property of the earthquake source models described above which is of central importance to the earthquake/underground explosion discrimination problem is that the earthquake source may be described as a shearing mechanism. This leads to the result that the radiation field may be characterized in terms of dipolar and higher order multipole contributions (Turnbull, 1976). Where this is the case, a significant portion of the radiated energy is propagated in the form of shear (SH and SV) body waves and Love and Rayleigh surface waves. In comparison, the explosive source, represented as a point dilatational source, generates only compressional wave radiation at the source.

The source mechanism corresponding to an underground explosion has been described by Sharpe (1946) and Carpenter (1967). Schematically, the explosion source mechanism may be described as follows: At the instant of detonation, material immediately surrounding the detonation point is vaporized. A shock front proceeds radially outward. Because the initial stresses associated with this shock wave greatly exceed the material strength of the containing medium, there is a zone of fracture where energy is dissipated in the work of nonelastic deformation. At some radial distance from the detonation point, sufficient energy has been dissipated for the laws of linear elastic

deformation to apply. At this point, energy is dissipated in the form of seismic radiation. The underground explosion source has been modelled in terms of an outward directed pressure pulse applied to the surface of a spherical cavity, which is equivalent to a radially symmetric point dilational source. In an infinite medium, the expected radiation field from a cavity source is purely compressional. In a layered halfspace, boundary effects result in some conversion of the P wave radiation to shear (SV) body waves and Rayleigh surface waves. However, Gilbert (1973), and Douglas, Hudson, and Kembhavi (1971) have shown that the partitioning of energy into the shear wave radiation field at the source results in a greater excitation of surface waves by the earthquake source. Moreover, discounting the effects of non-linear mode conversion, no SH shear body waves or Love surface waves are generated by the explosive source.

The theoretical seismic radiation patterns for a spherical cavity source assume a purely monopole distribution. However, a significant quadrupole contribution has been observed in the SH-wave and Love wave radiation fields for many explosions (Archambeau and Sammis, 1970; Archambeau, 1972). Archambeau (1972) suggested that this anomalous radiation field may be contributed by a combination of stress relaxation accompanying the creation of a fracture zone in a prestressed medium, and by weak zones and faulting in the immediate vicinity of the detonation point, leading to secondary ruptures with

definite fault symmetries. Radiation caused by the first mechanism is quadrupolar in nature, whereas that from the second exhibits dipolar and higher multipole contributions. This technique has been applied by Archambeau and Sammis (1970) to obtain source parameters from the Bilby underground explosion.

Discriminants Based on Earthquake/Underground Explosion Source Differences

Several studies have been made where the expected displacement amplitudes for the earthquake source have been compared with those from a buried explosive source. Douglas, Hudson and Kembhavi (1971) studied the relative excitation of body and surface waves for both earthquake and underground explosion point source models. They concluded that the earthquake source, represented as a point double couple, is a significantly more efficient generator of shear body wave and Rayleigh surface wave radiation than the explosive source, modelled as a point dilatational source. A similar result has been obtained by Gilbert (1970, 1973).

Numerous discriminants have been developed to differentiate between earthquakes and underground explosions. In addition to the depth, geographic, and first motion discriminants already mentioned, several discriminants have been proposed which are based upon

differences in the relative excitation of body and surface waves between the earthquake source and the explosion source.

A discriminant involving the difference between long-period surface wave magnitude (M_s) and short-period body wave magnitude (m_b) was originally suggested by Press, et al, (1963). Subsequent research has shown the M_s-m_b discriminant to be one of the most effective teleseismic discriminants available (Thirlaway, 1968; SIPRI, 1968; Basham, 1969; Liebermann and Pomeroy, 1969; Capon, et al, 1969; Evernden, 1969). For events recorded at teleseismic distances, M_s-m_b reliably separates earthquakes and underground explosions for m_b values as low as 5.0 (Thirlaway, 1968; SIPRI, 1968) or 4.75 (Evernden, 1969). Recent studies also show the need for regionalization of M_s-m_b to account for differing attenuation characteristics of various source-receiver paths (Marshall and Basham, 1972).

Frequency domain studies have been used to suggest several possible spectral discriminants (Aki, 1967; Brune, 1970; Wyss et al., 1971; Hasegawa, 1972; Alexander and Lambert, 1973). In essence, these discriminants reflect the greater amount of low frequency energy expected in earthquake spectra. This is a result of differences in source dimensions and source time histories between earthquakes and underground explosions.

Other discriminants have been suggested, such as those

based upon the greater complexity of the earthquake signature (SIPRI, 1968), and related discriminants suggested by autoregressive modelling studies applied to the earthquake and explosion signatures (Tjostheim, 1975).

The discriminants used for the analysis presented in this paper are chosen to reflect the theoretical differences in energy partitioning between the compressional and shear wave radiation fields for earthquakes and underground explosions. These discriminants consist of a set of energy ratios first studied by Booker and Mitronovas (1964), and estimated M_s/m_b ratios. M_s/m_b ratios were used instead of M_s-m_b differences because the M_s/m_b ratio is formally consistent with the Booker and Mitronovas energy ratio discriminants.

Assuming a linear relationship between M_s and m_b of the form $M_s=Am_b+B$, which is valid for a given range of m_b values, the M_s-m_b discriminant parameterizes the slope A upon m_b, whereas the M_s/m_b discriminant parameterizes the constant B upon m_b. Because the m_b values used to form the M_s/m_b discriminant in this study vary over a reasonably small range of m_b which is nearly the same for both populations, and because the slopes for both populations are nearly unity, the M_s/m_b discriminant yields results which do not differ significantly from those obtained using the M_s-m_b discriminant. F-statistics for both cases are given in Table (V-1).

CHAPTER III

MULTIVARIATE CLASSIFICATION TECHNIQUES

Once a suitable set of discriminants has been obtained, several methods may be used to incorporate the set into a multivariate classification scheme. Two techniques are discussed in this paper.

The first involves multiple discriminant analysis theory, where an event is classified according to the magnitude of a likelihood function derived from a Gaussian probabilistic model. Both linear and quadratic classification schemes are discussed.

The second technique is based on several cluster analysis procedures. The cluster analysis approach differs fundamentally from that of discriminant anlysis, in that no a priori probabilistic model is assumed. Instead, groups, or clusters, are allowed to form simply on the basis of the optimization of a specified similarity criterion. The immediate attraction of this non-parametric approach is offset, however, by the relatively poor performance of clustering techniques compared to that of the discriminant analysis procedures described below.

Multiple Discriminant Analysis: Theory

In this section, the problem of constructing a method for grouping or identifying similar events is discussed in terms of multiple discriminant analysis. This method incorporates a probabilistic interpretation to the problem, in that the event is considered to be a random observation from one of several different possible stochastic populations. The desired objective is to identify that population from which the observation originates. This is accomplished by constructing and testing the hypotheses, H_i, that the observation originates from the i^{th} population.

The hypothesis testing may be carried out according to one of several criteria of optimality. All of the criteria considered lead to a simple test involving comparisons of the magnitudes of the likelihood functions for each group. Moreover, where only two groups are in question, this procedure is further simplified to a binary likelihood ratio test.

The problem of multiple discriminant analysis was originally formulated by Fisher (1936), and has since been considered in some detail by Anderson (1958), Anderson and Badahur (1962), Whalen (1971), and Rao (1973). The treatment given here is derived principally from the work of these latter authors.

Classification Into One of Several Populations

Let $\vec{x}$ represent a p-component random vector of observations originating from one of q populations, $A_1,\ldots,A_q$. To classify a specific realization of $\vec{x}$, say $\vec{x}_o$, into one of these populations, the following q hypotheses are formed:

$$\begin{array}{lll} H_1: & \vec{x}_o \text{ originates from } A_1, \\ \vdots & \vdots & \vdots \\ H_q: & \vec{x}_o \text{ originates from } A_q. \end{array}$$

The problem now becomes one of testing the hypotheses, H_i, and identifying $\vec{x}_o$ with the population A_i with the largest probability that H_i is true. This is done according to some criterion which weighs, either directly or indirectly, the effects of incorrectly classifying the given observation.

The probability that the hypothesis H_i is true for a specific realization $\vec{x}_o$, of the random vector $\vec{x}$, may be found from Bayes' Rule. Let P represent the event that H_i is true, and let Q represent the event that $\vec{x} = \vec{x}_o$. Then the marginal, conditional, and joint probabilities of P ($\vec{x}$ originates from H_i) and Q ($\vec{x} = \vec{x}_o$) are given by:

$Pr\{P\} = Pr\{H_i\}$ = marginal probability that the hypothesis H_i is true,

$Pr\{Q\} = Pr\{\vec{x}=\vec{x}_o\}$ = marginal probability that $\vec{x} = \vec{x}_o$,

$Pr\{P|Q\} = Pr\{H_i|\vec{x}=\vec{x}_o\}$ = probability that H_i is true given that $\vec{x} = \vec{x}_o$,

$Pr\{Q|P\} = Pr\{\vec{x}=\vec{x}_o|H_i\}$ = probability that $\vec{x} = \vec{x}_o$, given that H_i is true,

$Pr\{P,Q\} = Pr\{Q,P\} = Pr\{H_i,\vec{x}=\vec{x}_o\} = Pr\{\vec{x}=\vec{x}_o,H_i\}$ = joint probability of $\vec{x} = \vec{x}_o$ and H_i.

Since the joint probability of P and Q is given by:

$$
\begin{aligned}
Pr\{P,Q\} &= Pr\{P/Q\}\, Pr\{Q\} \\
&= Pr\{Q/P\}\, Pr\{P\}, \qquad \text{(III - 1)}
\end{aligned}
$$

it is possible to obtain the following expression for $Pr\{P|Q\}$:

$$Pr\{P/Q\} = \frac{Pr\{Q/P\}\, Pr\{P\}}{Pr\{Q\}} \quad \text{(III - 2)}$$

(Bayes' Rule)

Pr{Q|P} and Pr{Q} may be rewritten in terms of the marginal and conditional probability density functions of Q. Recalling that Pr{Q|P} is the conditional probability that $\vec{x} = \vec{x}_o$, given that H_i is true, and that Pr{Q} is the marginal probability that $\vec{x} = \vec{x}_o$, the probability that $\vec{x}$ is located in a p-dimensional volume $dV = dx_1,\ldots,dx_p$ centered about $\vec{x}_o$ is approximated by:

$$Pr\{x_{o1} \le x_1 \le x_{o1} + dx_1, \ldots, x_{op} \le x_p \le x_{op} + dx_p\}$$
$$\cong p(\vec{x}_o)\, dV, \quad \text{(III - 3)}$$

where $p(\vec{x}_o)$ is the marginal probability density function of $\vec{x}$ evaluated at $\vec{x} = \vec{x}_o$. Similarly, the conditional probability that $\vec{x}$ is located in the volume dV, given that H_i is true, can be approximated by:

$$Pr\{[x_{o1} \le x_1 \le x_{o1} + dx_1, \ldots, x_{op} \le x_p \le x_{op} + dx_p]/H_i\}$$
$$\cong p_i(\vec{x}_o)\, dV \quad \text{(III - 4)}$$

where $p_i(\vec{x}_o) = p_i(\vec{x}_o|H_i)$ is the conditional probability density function of x, given that H_i is true, evaluated at $\vec{x} = \vec{x}_o$. For $\vec{x}$ continuous in dV, as dV becomes small, the approximations above become exact. Also, because $p(\vec{x}_o)$ and

$p_i(\vec{x}_o)$ are constant for a given value of $\vec{x}_o$, the probabilities in (III-3) and (III-4) approach zero with dV. However, in the ratio given by Bayes' Rule, dV is present in both the numerator and denominator, so that it cancels. Where this is the case, the probability that H_i is true, given a realization, $\vec{x}_o$, of the random vector $\vec{x}$, may be written:

$$Pr\{H_i / \vec{x} = \vec{x}_o\} = \frac{p_i(\vec{x}_o)\, Pr\{H_i\}}{p(\vec{x}_o)} \qquad \text{(III - 5)}$$

In many cases it is sufficient to specify that $\vec{x}$ be classified into the population which maximizes $Pr\{H_i | \vec{x} = \vec{x}_o\}$. This is actually a special case of the Bayes' criterion, where the assigned costs of misclassifying an observation are equal, and the costs of correctly classifying the observation are zero. On the other hand, it often happens that this rule cannot be used. This is true, for example, in cases where the prior probabilities of the hypotheses ($Pr\{H_i\}$) are unknown, or where the experimenter wishes to weight the relative importance of various kinds of misclassification. It is desireable, then, to have classification criteria available which allow for both of these instances.

Three general criteria are presented in Appendix A. The Bayes' criterion is the most powerful, in that it allows for specification of weights (costs) on the various types of

misclassification. However, the Bayes' criterion requires that the exact prior probabilities for the hypotheses be specified. The minimax criterion allows for the specification of costs, but does not require the specification of prior probabilities. The Neymann-Pearson criterion specifies the acceptable threshold probability for misclassification of $\vec{x}_o$ into a particular population, and requires neither costs nor explicit specification of prior probabilities.

Classification Into One of Two Multivariate Normal Populations

The problem considered in this section is that of classifying a seismic event, based upon multiple measurements at each station, averaged over the total number of stations. This is the problem originally considered by Booker and Mitronovas (1964). Specifically, the two populations consist of earthquakes and underground nuclear explosions. The measurement random variables are the energy ratios of Booker and Mitronovas (1964), augmented with estimated M_s/m_b ratios.

If $\vec{x}$ represents the random vector of these measurements, $\vec{x}$ is assumed to have the multivariate normal distribution $N(\vec{m}_1, C_1)$ if $\vec{x}$ originates in the earthquake population, and $N(\vec{m}_2, C_2)$ if $\vec{x}$ originates in the underground

explosion population. Here $\vec{m}_1$, C_1, and $\vec{m}_2$, C_2 are the mean vectors and covariance matrices of the earthquake and explosion populations, respectively. These distributions are given in terms of the likelihood functions $p_1(\vec{x})$ and $p_2(\vec{x})$:

$$p_1(\vec{x}) = \frac{1}{(2\pi)^{p/2}|C_1|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{m}_1)' C_1^{-1}(\vec{x}-\vec{m}_1)\right\} \qquad \text{(III - 6)}$$

and

$$p_2(\vec{x}) = \frac{1}{(2\pi)^{p/2}|C_2|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{m}_2)' C_2^{-1}(\vec{x}-\vec{m}_2)\right\} . \qquad \text{(III - 7)}$$

Defining the hypotheses H_1 and H_2 for the specific realization $\vec{x} = \vec{x}_o$:

H_1: $\vec{x}_o$ originates in A_1,

and

H_2: $\vec{x}_o$ originates in A_2,

it is possible to construct the likelihood ratio test for H_1:

Choose H_1 if $p_1(\vec{x}_o)/p_2(\vec{x}_o) \geq \lambda_T$,

and choose H_2 otherwise,

where λ_T is the threshold value determined from the classification criterion (see Appendix A).

Substituting the normal densities given for $p_1(x_o)$ and $p_2(x_o)$ in (III-6) and (III-7), the likelihood ratio becomes:

$$\frac{p_1(\vec{x})}{p_2(x)} = \frac{|C_2|^{\frac{1}{2}}}{|C_1|^{\frac{1}{2}}} \exp\left\{-\tfrac{1}{2}\left[(\vec{x}_o-\vec{m}_1)'C_1^{-1}(\vec{x}_o-\vec{m}_1) - (\vec{x}_o-\vec{m}_2)'C_2^{-1}(\vec{x}_o-\vec{m}_2)\right]\right\}. \qquad \text{(III - 8)}$$

Because the logarithm is a monotonically increasing function, the likelihood ratio test for H_1 may be rewritten in terms of the log-likelihood ratio test without changing the direction of the inequality:

Choose H_1 if

$$-\tfrac{1}{2}\left[(\vec{x}_o-\vec{m}_1)'C_1^{-1}(\vec{x}_o-\vec{m}_1) - (\vec{x}_o-\vec{m}_2)'C_2^{-1}(\vec{x}_o-\vec{m}_2)\right]$$

$$\geq \log \lambda_T + \tfrac{1}{2}\log\frac{|C_1|}{|C_2|}. \qquad \text{(III - 9)}$$

If the covariance matrices, C_1 and C_2 are equal, the expression to the left of the inequality is linear in $\vec{x}$. However, if $C_1 \neq C_2$, the expression on the left is quadratic in $\vec{x}$. In their analysis, Booker and Mitronovas (1964) made the assumption that the covariance matrices are equal, thereby linearizing the likelihood ratio. Where this is the case, the likelihood ratio test becomes:

Choose H_1 if

$$\vec{x}_0' C^{-1}(\vec{m}_1 - \vec{m}_2) + \frac{1}{2}(\vec{m}_2 - \vec{m}_1)' C^{-1}(\vec{m}_2 + \vec{m}_1) \geq \log \lambda_T \qquad \text{(III - 10)}$$

The log-likelihood ratio on the left of (III-10) is known as Fisher's (1936) discriminant function. By letting

$$\vec{F} = C^{-1}(\vec{m}_1 - \vec{m}_2) \qquad \text{(III - 11)}$$

and

$$k = \frac{1}{2}(\vec{m}_2 - \vec{m}_1)' C^{-1}(\vec{m}_2 + \vec{m}_1), \qquad \text{(III - 12)}$$

the Fisher discriminant function may be written:

$$U(\vec{x}_0) = \vec{x}_0' \vec{F} + k \qquad \text{(III - 13)}$$

In this final form, the linear log-likelihood ratio test for H_1 is given by:

Choose H_1 if $U(\vec{x}) \geq \log(\lambda_T)$,
choose H_2 otherwise.

For $\vec{x} = \vec{x}_T$, where $\vec{x}_T$ is chosen so that

$$U(\vec{x}_T) = \vec{x}_T' \vec{F} + k = \log \lambda_T, \qquad \text{(III - 14)}$$

the function $U(\vec{x}_T)$ describes the surface which optimally divides the parameter space of $\vec{x}$ into the two classification regions R_1 and R_2. Because $U(\vec{x})$ is linear in $\vec{x}$, $U(\vec{x}_T)$ defines a hyperplane in the parameter space $R = (R_1 \cup R_2)$. This situation is illustrated schematically for the univariate case in Figure (III-1).

On the other hand, if the two population covariance matrices are not equal, the surface dividing R is quadratic in $\vec{x}$:

$$\begin{aligned} U(\vec{x}_T) &= \tfrac{1}{2}\vec{x}_T'(C_2^{-1} - C_1^{-1})\vec{x}_T + \vec{x}_T'(C_1^{-1}\vec{m}_1 - C_2^{-1}\vec{m}_2) \\ &\quad + \tfrac{1}{2}(\vec{m}_2' C_2^{-1}\vec{m}_2 - \vec{m}_1' C_1^{-1}\vec{m}_1) \qquad \text{(III - 15)} \\ &= \log \lambda_T + \tfrac{1}{2}\log\frac{|C_1|}{|C_2|} . \end{aligned}$$

In the two-dimensional case, $U(\vec{x})$ describes a quadratic curve, either an ellipse or a pair of hyperbolas, where the decision regions R_1 and R_2 are either the areas interior and exterior to the ellipse, or the areas bounded by the pair of hyperbolas (Anderson and Badahur, 1962). A likelihood ratio test based on a quadratic log-likelihood ratio relies quite heavily on the assumption of normality, because the classification is determined to a great extent by those values of the likelihood functions located at some distance from the mean where the probability densities are small.

This situation is illustrated for the univariate case in Figure (III-2).

The case where the population covariance matrices are unequal and the assumption of normality is valid for regions close to the mean, but not necessarily for regions at any great distance from the mean, has been considered by Anderson and Badahur (1962). Their analysis consists of finding the linear function (hyperplane) which minimizes a function of the error probabilities that corresponds to either the Neymann-Pearson, Minimax, or Bayes′ criterion. If the linear function U^q is given by:

$$U^q(\vec{x}) = \vec{x}'\vec{r} + \text{const.} \qquad \text{(III - 16)}$$

the vector $\vec{r}$ is found to be the product of the inverse of a weighted sum of the covariance matrices and the difference between the population means:

$$\vec{r} = (t_1 C_1 + t_2 C_2)^{-1}(\vec{m}_1 - \vec{m}_2) \, . \qquad \text{(III - 17)}$$

Anderson and Badahur (1962) suggest solving for $\vec{r}$ by iteratively substituting for t_1 and t_2 until an auxillary condition (determined by the function of the error probabilities used as an optimality criterion) is satisfied. This technique has the advantage of being more robust to deviations from the normal assumption than the quadratic procedure discussed previously; however, it is far less

efficient computationally than the techniques described for the linear and quadratic procedures above.

Classification based on Samples taken from the Respective Populations

In many problems, including that of classifying earthquakes and underground explosions, the parameters (i.e. the mean vectors and covariance matrices) of the respective populations are not known a priori, and must be estimated from a "training set" of data. Letting $\vec{x}^1{}_1,\ldots,\vec{x}^1{}_{N1}$ represent a set of observations from A_1 and $\vec{x}^2{}_2,\ldots,\vec{x}^2{}_{N2}$ represent a similar set from A_2, the problem becomes one of obtaining the discriminant function $U(\vec{x})$ so that a new observation ($\vec{x}_o$) may be classified into A_1 or A_2.

In the linear problem where the population covariance matrices are assumed to be equal, the sample mean vectors for the two populations are given by:

$$\bar{x}_1 = \frac{1}{N_1}\sum_{j=1}^{N_1} \vec{x}_j^1$$
$$\bar{x}_2 = \frac{1}{N_2}\sum_{j=1}^{N_2} x_j^2 \quad \text{(III - 18)}$$

The estimate for the common covariance matrix is given by the within groups sample covariance matrix:

$$(N_1 + N_2 - 2) S = \sum_{i=1}^{2} \left[\sum_{j=1}^{N_i} (x_j^i - \bar{x}_i)(x_j^i - \bar{x}_i)' \right] \qquad \text{(III - 19)}$$

Substitution of $\vec{\bar{x}}_1$, $\vec{\bar{x}}_2$, and S into the log likelihood ratio in (III-10) gives the Fisher (1936) discriminant function for the case where the parameters are estimated:

$$U(\vec{x}) = \vec{x}' S^{-1} (\bar{x}_1 - \bar{x}_2) - \tfrac{1}{2} (\bar{x}_1 + \bar{x}_2) S^{-1} (\bar{x}_1 - \bar{x}_2) \qquad \text{(III - 20)}$$

By applying the discriminant function given in (III-20) to the training set, it is possible to evaluate how well the discriminant analysis scheme separates the sample observation vectors. Estimates of the overall average probability of correct classification may be obtained by evaluating $\Pr\{H_i | \vec{x} = \vec{x}_o\}$ for each observation in the training set. In terms of the likelihood functions $p_1(\vec{x})$ and $p_2(\vec{x})$, the probability that H_i is true given that $\vec{x} = \vec{x}_o$ is defined in (III-5) as:

$$\Pr\{H_i / \vec{x} = \vec{x}_o\} = \frac{p_i(x_o) \Pr\{H_i\}}{p(x_o)}, \qquad \text{(III - 21)}$$

where $p(\vec{x})$ is the marginal probability density of $\vec{x}$, given by:

$$p(\vec{x}) = \sum_{i=1}^{2} p_i(\vec{x}) \Pr\{H_i\} \qquad \text{(III - 22)}$$

For a specified set of prior probabilities, (III-21) may be evaluated for each observation vector in the training

set to give the classification probabilities for each event. The average probability of correct classification for the training set may then be found by averaging the classification probabilities for each event over the two populations. Inasmuch as the parent population for each event in the training set is known, if $\vec{x}_o$ denotes a sample vector from A_1, then $Pr\{H_1|\vec{x}=\vec{x}_o\}$ is the probability of correctly classifying $\vec{x}_o$ and $Pr\{H_2|\vec{x}=\vec{x}_o\}$ is the probability of misclassifying $\vec{x}_o$. The overall probability of correct classification in terms of the samples in the training set is given by:

$$P_c = \frac{1}{N_1}\sum_{j=1}^{N_1} P_A\{H_1/\vec{x}=\vec{x}_j^{\,1}\} + \frac{1}{N_2}\sum_{j=1}^{N_2} P_A\{H_2/\vec{x}=\vec{x}_j^{\,2}\}. \qquad \text{(III - 23)}$$

In a like fashion, the overall average probability of misclassification is given by:

$$P_e = \frac{1}{N_1}\sum_{j=1}^{N_1} P_A\{H_2/\vec{x}=\vec{x}_j^{\,1}\} + \frac{1}{N_2}\sum_{j=1}^{N_2} P_A\{H_1/\vec{x}=\vec{x}_j^{\,2}\}. \qquad \text{(III - 24)}$$

In addition to the determination of the classification probabilities P_c and P_e, a discriminant analysis of the training set may provide an insight into which subset of variables best contributes to the separation of the two

populations. The Mahalanobis D^2 statistic may be used as a measure of the separation of the two populations (Anderson, 1958). The Mahalanobis D^2 statistic is defined (Rao, 1973):

$$D^2 = (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2) \tag{III - 25}$$

This statistic may be normalized and used as an F-statistic to test the hypothesis that the two population means are equal. The D^2 statistic (or its associated F-statistic) may be calculated for any single variable in the discriminant analysis, or for any group of variables. This provides a tool for determining which of the measured variables contributes the greatest amount of information to the discriminant analysis procedure.

Cluster Analysis Techniques

In addition to the multiple discriminant analysis theory presented above, several cluster analysis techniques are applied to the Booker and Mitronovas (1964) energy ratios and the estimated M_s/m_b ratios. Cluster analysis appears to be an attractive means for grouping data because no external probabilistic model need be assumed. Instead, cluster groupings are formed according to the optimization of a specified criterion. For example, in this study the criterion used to form clusters is the minimization of the

total error sum of squares.

A cluster may be viewed as a collection, or set of events, where the total error sum of squares is minimized. In equivalent terms, the Euclidean distance from each event to the centroid of the parent cluster is less than the distance to the centroid of any other cluster. In the case of the Booker and Mitronovas data set, each event is considered as a point in a seven-dimensional space (or an eight-dimensional space where estimated M_s/m_b ratios are included). The similarity measure used to fuse these events into clusters is the squared Euclidean distance. The distance between the i^{th} and j^{th} event in a p-dimensional space is defined as:

$$d_{ij}^2 = \sum_{k=1}^{p} (x_{ik} - x_{jk})^2. \qquad \text{(III - 26)}$$

The clustering techniques used in this study consist of Ward's hierarchical clustering procedure and an iterative relocation procedure, both of which are described by Everitt (1974) and Anderberg (1973).

The strategy used in Ward's procedure involves first constructing a similarity matrix containing the squared Euclidean distance between each pair of events in the input data set. In the beginning, each event is viewed as a single cluster. Events are fused in such a way that the smallest increase in the error sum of squares occurs. The error sum of squares is defined as the sum of the distances from each

event to the centroid of its parent cluster (Wishart, 1975), where the cluster centroid is simply the cluster's center of mass. At the first step, with each event viewed as a single cluster, the total error sum of squares is zero. The two closest events are fused to form a single cluster. This process proceeds, either by fusing pairs of events, by adding events to existing clusters of events, or by fusing clusters until only one large cluster remains.

Ward's method has the disadvantage that cluster groupings obtained at an early level in the procedure remain fixed throughout the clustering process. That is, if two events are fused early in the clustering procedure, they remain as members of the same cluster throughout, even though at some later point a decrease in the total error sum of squares might be realized by assigning the events to different clusters. This problem can be avoided by using an iterative relocation procedure, such as the one described by Wishart (1975).

In Wishart's RELOCATE procedure, at any given step, each object is considered in turn, and removed from its parent cluster. The distance from the removed event to the centroid of each cluster is computed, and the event is returned to that cluster for which this distance is minimized. After the relocation cycle has been completed (that is, when no more events can be moved), the two most similar clusters are fused. These relocation and fusion procedures may then be alternated until the desired number

of clusters is obtained.

The clustering procedures described in this section and the multiple discriminant analysis procedures described in the previous section were both used to classify a set of earthquakes and underground explosions originally studied by Booker and Mitronovas (1964). The next chapter contains a description of the Booker and Mitronovas data set, the programs used in the classification schemes, and the results obtained by applying these programs to the Booker and Mitronovas data set.

CHAPTER IV

APPLICATION OF CLUSTERING AND DISCRIMINANT ANALYSIS TECHNIQUES TO A SET OF EARTHQUAKES AND UNDERGROUND EXPLOSIONS

Description of the Data

The data analyzed using the discriminant analysis and clustering techniques described earlier consist of a set of 20 earthquakes and 27 Nevada Test Site (NTS) underground nuclear explosions. Summary information for these events is provided in Table (IV-1). With the exception of three explosion collapses which have been deleted in the present study, this is the same set of events studied by Booker and Mitronovas (1964), and by Mitronovas (1963).

The measurement variables on which the discriminant and cluster analyses are based are the set of seven energy ratios published by Booker and Mitronovas (1964). In addition, for events where the body wave magnitudes have been calculated, the original set of variables is augmented with estimated M_s/m_b ratios, where the M_s estimates were randomly generated about the M_s/m_b lines. Measured M_s values could not be used because the long-period data for these events are either nonexistent or of such poor quality

that reliable M_s estimates could not be obtained. Instead, the M_s estimates were generated by drawing upon recent existing results relating M_s to m_b for NTS explosions and earthquakes. The M_s/m_b ratios used were calculated by generating values for M_s, normally distributed with a specified variance, about the expected point on the appropriate M_s/m_b line. The M_s/m_b lines used in this study are those given by Alexander and Lambert (1973), typical of underground explosions and small-magnitude earthquakes in the Western United States. Four M_s/m_b ratios were calculated for each event using published body wave magnitudes. In each case the variance used in generating the earthquake M_s estimates is twice that used in generating the M_s estimates for the underground nuclear explosions. This greater uncertainty is incorporated to reflect the assumed complexity and variability of the earthquake source mechanism, as compared to that of the underground explosion source mechanism. The explosion M_s variance estimates were chosen as multiples of the M_s variance given by von Seggern (1972). Table (IV-2) lists the energy ratios used in this study. The m_b values, the M_s estimates and the M_s/m_b ratios are given in Table (IV-3). M_s versus m_b is plotted for each variance level in Figures (IV-1) through (IV-4).

The energy ratios for each station were originally obtained by Booker and Mitronovas by calculating the energies within time intervals corresponding to specified velocity windows for each event, and forming ratios for

those window combinations which appeared to give promising discrimination criteria. The total energy within each window for a given event was found by summing over the square of the vertical, radial, and transverse components at each station, then averaging over the stations. These values were corrected for microseismic noise by subtracting the average noise energy based on samples taken in a window immediately preceding the first arrival for each event.

The above calculations were performed on 302 seismograms representing the 50 events. Each seismogram was recorded on Long Range Seismic Monitoring (LRSM) network short-period Benioff seismometers. Because normally the LRSM horizontal radial component is oriented toward NTS, the horizontal components were appropriately rotated prior to evaluating the energy ratios for each earthquake. The velocity windows used in forming the energy ratios are listed in Table (IV-4).

Discriminant Analyses

The discriminant analysis in this paper consists of two parts. The first part is simply an attempt to duplicate the results obtained by Booker and Mitronovas (1964) using the 1975 revised version of BMD04M, a two group linear multiple discriminant analysis program supplied in the Biomedical Computer Programs Package (BMD), and published by the UCLA

School of Medicine. These results are compared with those obtained using a more recent linear multiple discriminant analysis program, BMD07M, supplied in the same package, and with the results obtained using the quadratic discriminant analysis program described in Appendix B.

The second part consists of a discriminant analysis of the subset of the Booker and Mitronovas events which has been augmented with the estimated M_s/m_b ratios described above. The M_s/m_b discriminant is one of the most successful discriminants yet found, so that its introduction should greatly enhance the performance of the original Booker and Mitronovas discriminant analysis scheme.

Program Descriptions

The BMD04M program calculates the linear term of the Fisher (1936) discriminant function for each event. If the i^{th} event is denoted by $\vec{x}_i$, this linear term is given by:

$$u(\vec{x}_i) = \vec{x}_i{}' S^{-1} (\bar{x}_1 - \bar{x}_2), \qquad \text{(IV - 1)}$$

where $\vec{x}_1$ and $\vec{x}_2$ denote the respective sample means vectors of the earthquake and explosion populations, and S is the within-group scatter matrix, given by:

$$S = \sum_{\alpha=1}^{2} \left[(\vec{x}_j - \bar{x}_\alpha)(\vec{x}_j - \bar{x}_\alpha)' \right]. \qquad \text{(IV - 2)}$$

The values of $U(\vec{x}_i)$ are then listed in order of decreasing algebraic magnitude. Because the constant term of the discriminant function is common to all events, and does not affect the ordering, it is not included in the calculation of the $U(\vec{x}_i)$.

In addition to computing and rank ordering the $U(\vec{x}_i)$, BMD04M computes the Mahalanobis D^2 statistic, given by:

$$D^2 = (\bar{x}_1 - \bar{x}_2)' S^{-1} (\bar{x}_1 - \bar{x}_2) , \qquad \text{(IV - 3)}$$

and the associated F-statistic for testing the differences in group means:

$$F(p, N_1 + N_2 - p - 1) = \frac{N_1 N_2 (N_1 + N_2 - p - 1)}{p (N_1 + N_2)(N_1 + N_2 - 2)} D^2. \qquad \text{(IV - 4)}$$

Program BMD07M also computes the linear discriminant function. However, BMD07M is more flexible than BMD04M in that provision is made for the introduction of prior probabilities, whereas BMD04M assumes that all prior probabilities are equal. BMD07M is able to classify into more than two groups, and calculates posterior classification probabilities for each event. However, the most useful characterisrtic of this program is the way in which variables are entered into the discrimination process.

BMD04M calculates initial F-statistics for each variable alone, then constructs classification functions using the entire specified set of variables. BMD07M, on the other hand, computes the initial F-statistics for each variable, then enters variables in a stepwise fashion, where the F-to-enter and F-to-remove statistics are updated at each step. This feature allows for a comparison of the relative discriminating power among variables, as well as for each variable taken alone. For example, two variables may have initial F-statistics of comparable magnitude. Upon entering the variable with the larger initial F value, the F-to-enter for the second variable may decrease significantly. This indicates that the two variables are highly correlated, so that the second variable contributes only a marginal amount of additional information to the discrimination process. Because this facility allows the program to be used to obtain maximally independent sets of discriminant variables, BMD07M should prove to be particularly valuable in studying the earthquake/underground explosion discrimination problem.

Quadratic discriminant procedures provide a means for discriminating between groups where the events are normally distributed, but where the covariance matrices differ markedly. Such procedures sacrifice some of the robustness of the linear method to deviations from the normal hypothesis, but they may provide better classifications than the linear procedures where the normal hypothesis is applicable and the covariance matrices for the two groups

differ significantly. A description of the quadratic discriminant analysis program used in this study is provided in Appendix B.

Discriminant Analysis Results for the Original Booker and Mitronovas Data Set

The classification obtained using BMD04M is very similar to the one obtained by Booker and Mitronovas (1964), except that several adjacent events have been interchanged in the rank ordering. The misclassification of the Boxelder Creek earthquake occurs in the same fashion here as in the previous studies. The classification obtained using BMD07M where equal prior probabilities are assumed is identical to that obtained using BMD04M. This assumption of equal prior probabilities is made for all the discriminant analyses presented here. The ratios, the order in which they are entered, the initial F-statistics, and the F-to-enter and F-to-remove values at each step are given in Table (IV-5).

Before any of the variables are entered, the value of the F-statistic based on the Mahalanobis D^2 statistic is calculated for each variable. This is a measure of the separation in group means achieved by each variable individually. As shown in Table (IV-5), variable 4 (Pg/Lg_1) is the most powerful discriminant among the Booker and Mitronovas energy ratios. Next is variable 3 (Pg/B), then

variable 1 (P_1/S_1), variable 5 ($Pg/[Lg_1+Rg_1]$), and variable 2 (P_2/S_2). Variables 6 (Rg_1/Lg_1) and 7 (Rg_2/Lg_2), where taken alone, do not appear to be promising discriminants.

The first variable to be entered is variable 4. Of the remaining 6 variables, variables 2 and 5 provide the greatest additional discriminating power. Upon entering variable 2 at the second step, variables 1 and 3 become the most powerful of those yet to be entered. This indicates a strong correlation between variables 2 and 5, so that variable 5 contributes only marginally to the discriminant process after variable 2 has been entered. The process of entering the variables continues until either all variables have been entered, or until none of the remaining variables can meet the F-to-enter threshold. The F-to-enter threshold is set by default to a small nominal value (0.001). However, this value may be adjusted at the user's option.

Using only the original Booker and Mitronovas energy ratios, the probability of correctly classifying an earthquake is 0.821. That of correctly classifying an explosion is 0.771. The overall average probability of correct classification is 0.792. The two events that are misclassified are the Boxelder Creek earthquake, identified as an explosion with probability 0.67, and the Codsaw explosion, marginally classified as an earthquake with probability 0.58.

BMD07M also computes the canonical transformation for the set of events under consideration. The canonical

variables are determined from a linear transformation of the original set of variables which maximizes the between-group scatter with respect to the within-group scatter for a given partitioning of the data set. Separation of the group means are reflected in the first canonical variable in the two-group case, where the remaining variables reflect within-group differences. A scatterplot of the first canonical variable versus the second is shown for the entire set of 27 explosions and 20 earthquakes in Figure (IV-5).

The final discriminant analysis of the original Booker and Mitronovas data set given here consists of a quadratic discriminant analysis using the program described in Appendix B. The Boxelder Creek earthquake is misclassified under this quadratic scheme as it was under the linear scheme, but with probability 0.95 instead of 0.67. However, instead of marginally misclassifying the Codsaw underground explosion as an earthquake as happens using the linear procedure, the quadratic procedure correctly identifies this explosion, but marginally identifies the Cache Creek earthquake as an explosion with probability 0.53. In both the linear and quadratic cases, however, the marginal misclassifications may be avoided by a judicious choice of the prior probabilities. Under the quadratic procedure, the average posterior probability of correctly classifying either an earthquake or an underground explosion is 0.94.

Discriminant Analysis Results for the Reduced Data Set with Estimated M_s/m_b Ratios

As mentioned in the data description above, a subset of the events used by Booker and Mitronovas have associated with them published body wave magnitudes. This subset consists of the nine earthquakes and 21 underground explosions shown in Table (IV-3). Because published M_s values are not available for any of these events, estimated M_s/m_b ratios have been used, where the M_s values are typical (statistically) of small magnitude earthquakes and underground explosions in the Western United States.

The first step in this classification consists of a linear discriminant analysis where BMD07M is applied to this subset of events, but where only the seven Booker and Mitronovas energy ratios are used. Under this linear procedure, where equal prior probabilities are assumed, eight of the earthquakes are correctly classified and one earthquake (Cache Creek) is misclassified as an explosion with probability 0.53. The Boxelder Creek event is not included in this analysis, because published m_b values for this event could not be found. Of the explosions, 19 are correctly classified and two (Codsaw and Stoat) are misclassified as earthquakes with probabilities 0.65 and 0.85 respectively. The average posterior probability of correctly classifying an earthquake is 0.81, and that of correctly classifying an underground explosion is 0.77. The

overall average probability of correct classification is 0.78. A summary of the stepwise discrimination process for this set of events is given in Table (IV-6), and a scatterplot of the first two canonical variables is given in Figure (IV-6).

The initial F-statistics for this reduced set of events is comparable to that found for the entire set of 20 earthquakes and 27 explosions. Taken alone, variable 4 (Pg/Lg_1) gives the greatest separation of the group means. Variables 3 (Pg/B), 1 (P_1/S_1) and 5 ($Pg/[Lg_1+Rg_1]$) are next and exhibit comparable discriminating potential. Variables 6 (Rg_1/Lg_1) and 7 (Rg_2/Lg_2) exhibit an insignificant separation of group means.

The quadratic discriminant analysis procedure, when applied to this subset of events, yields excellent results with no events misclassified. The average posterior probability of correctly classifying an earthquake is 0.95, and that of correctly classifying an explosion is 0.99, which yields an overall average probability of correct classification of 0.98.

It should be noted however, in the case of the quadratic discriminant procedure, that the results obtained for this reduced set of events might be somewhat misleading. The reason for this is that the covariance matrices for a given population are calculated only from the events in that population. Because the earthquake population in the reduced set consists of only nine events, with measurements

on seven variables (or eight variables where the M_s/m_b ratio is included), a reliable estimate for the population covariance matrix may not have been obtained. This is not necessarily the case for the linear procedure, where the single common covariance matrix is calculated as a weighted average of the individual population covariance matrices, where the weights are determined from the number of events in each group.

The effect of adding the M_s/m_b ratios to the original set of energy ratios is quite dramatic. A summary of the stepwise procedure used in BMD07M is listed in Tables (IV-7) to (IV-10) for each of the four sets of M_s/m_b ratios, where M_s/m_b is listed as variable 8. Scatterplots of the first two canonical variables for each set of M_s/m_b ratios are given in Figures (IV-7) to (IV-10).

In each case the M_s/m_b ratio was found to be the most powerful discriminant of the eight variables employed and, in each case, was the first variable entered. For the first two sets, where the variance of the explosion M_s estimates is equal to and 1.5 times the variance found by von Seggern (1972), the M_s/m_b ratio alone is sufficient to discriminate successfully between the two groups. For the third and fourth sets, where the explosion M_s variance is twice and three times times that given by von Seggern, only one misclassification occurs where the M_s/m_b ratio is used alone. For the third set, this misclassification is

corrected by entering the remaining seven variables (energy ratios). However, an interesting situation occurs for the fourth set, as shown in Table (IV-10). The first variable entered is variable 8 (M_s/m_b). At this point, the Passaic explosion is misclassified as an earthquake. The next variable entered is variable 3 (Pg/B), and at this step Passaic is correctly classified as an explosion. Variable 2 (P_2/S_2) is entered next, and the classification remains unchanged with no events misclassified. However, upon entering variable 5 ($Pg/[Lg_1+Rg_1]$), Passaic is again misclassified as an earthquake and remains so throughout the rest of the procedure. The Passaic explosion was not misclassified in any of the other discriminant analyses, and is only marginally misclassified here with probability 0.60. From Tables (IV-3) and (IV-10), and Figure (IV-4) it can be seen that the estimated M_s/m_b ratio for the Passaic explosion is more than two standard deviations from the explosion mean M_s value, and is within one standard deviation of the mean M_s value for earthquakes. For this reason, misclassification of the Passaic explosion may be expected where the M_s/m_b ratio is used alone. Incorporation of the other energy ratios shifts the event back and forth between the two populations, resulting in a final marginal misclassification. Because the classification probability for the Passaic explosion is closer to 0.5 than that of any other event identified as an earthquake, the Passaic explosion may be correctly classified by using a different

choice for the prior probabilities.

Where the quadratic discriminant analysis procedure was used, no events were misclassified. For the energy ratios augmented with the first three sets of M_s/m_b ratios, the average probability of correct classification for earthquakes and explosions is nearly 1.0. For M_s/m_b set 4, the probability of correctly classifying an explosion is nearly 1.0, and the probability of correctly classifying an earthquake is 0.99.

A summary of the classification probabilities for each of the discriminant analyses is given in Table (IV-11).

Cluster Analysis Results

Program Descriptions

The cluster analysis procedures used in this paper are supplied in the Clustan IC cluster analysis package developed at University College, London (Wishart, 1975). As mentioned in Chapter III, two clustering techniques are employed in this study. The first is Ward's hierarchical clustering procedure. This method involves constructing a similarity matrix containing the squared Euclidean distance between each pair of events in the input data set. At the start of the procedure, each event is considered as a single

cluster, and the two events with the smallest Euclidean distance between them are fused to form a new cluster. A new cluster centroid (center of mass) is computed, the total error sum of squares is calculated, and the similarity matrix is updated where the centroid of the new cluster is substitued for the old events or cluster centroids. In this way, the rank of the similarity matrix decreases by one each time a new cluster is formed. This process continues until the desired number of clusters has been obtained. At each step, the events, or clusters, which are fused are those which result in the smallest increase in the total error sum of squares.

In the case of the hierarchical procedure contained in the Clustan IC package, the cluster fusions are effected by a combinatorial transformation of the distance matrix. Assuming that the clusters P and Q have just been fused, the increase in the error sum of squares which occurs upon fusion of the new cluster (P+Q) with any other cluster R, is given by (Wishart, 1975):

$$S(R, P+Q) = \frac{1}{N_R + N_P + N_Q} \Big[(N_R + N_P)\, S(R,P) + (N_R + N_Q)\, S(R,Q) - N_R\, S(P,Q) \Big]. \qquad \text{(IV - 5)}$$

where N_p, N_q, and N_r are the respective sizes of the clusters P, Q, and R.

The hierarchical clustering method has the disadvantage

that cluster groupings obtained at an early point in the clustering procedure must remain fixed throughout the process. In other words, if two events are fused at some point to form a new cluster, these two events must remain members of the same cluster throughout the clustering process. This is the case even where a net decrease in the total error sum of squares might be obtained if the two events were assigned to different clusters. The RELOCATE procedure in the Clustan IC package circumvents this problem. The RELOCATE procedure uses a convergent K-means method (Everitt, 1974) to relocate events iteratively in new clusters at each step in the hierarchical procedure. At any given step, each event is removed from its parent cluster, and the Euclidean distances between the event and the centroids of every cluster are computed. If the distance from the event to any other cluster centroid is less than the distance from the event to the centroid of the parent cluster, the event is reassigned to the new cluster. The two closest clusters are then fused and the entire process is repeated until the desired number of clusters has been obtained. This method has the additional advantage that a similarity matrix need not be stored in core or on disk. On the other hand, this method is several times slower than the hierarchical procedure.

Cluster Analysis Results

Because the cluster analysis procedures used in this study seek to minimize the total error sum of squares, inclusion of variables which do not reflect given group differences can mask the contributions from those which do reflect such differences. For this reason, the cluster analyses described here were based on the three most powerful discriminant variables obtained from the discriminant analyses. In this way the tendency for the data to cluster into the actual earthquake and explosion populations is enhanced.

Application of Ward's hierarchical method and the K-means iterative relocation procedure to the original set of 20 earthquakes and 27 underground explosions shows that the resulting clusters do not reflect actual group divisions. A fusion tree which shows the cluster fusion history from Ward's method is given in Figure (IV-11). This analysis was generated using variables 4 (Pg/Lg_1), 2 (P_2/S_2), and 3 (Pg/B).

It has been suggested by Wishart (1975) that stable clusters are those which require a large increase in the total error sum of squares in order to fuse them with other clusters. With this in mind, it may be seen from Figure (IV-11) that the cluster analysis of the Booker and Mitronovas data set is stable at the four-cluster level. A move to the three-cluster level results in an increase in

the total error sum of squares by more than 100 per cent.

Scatterplots are given in Figures (IV-12) and (IV-13) which show the cluster distributions at the four-cluster level. In both figures, variable 4 (Pg/Lg_1) is plotted versus variable 2 (P_2/S_2). Figure (IV-12) was obtained from the cluster groupings generated by Ward's method, and Figure (IV-13) was obtained from those generated by RELOCATE, beginning with a random partition into five initial clusters. The cluster memberships are identical in both cases, although the indexing for the clusters varies because of the internal book-keeping of each procedure. The close similarities between these two groupings indicate an optimal grouping because each was obtained in a different manner. However, in both cases, no single cluster consists exclusively of earthquakes or explosions.

The next step consists of cluster analyses similar to those described above for the reduced set of events, both in the case where only the seven energy ratios are used, and where the lowest variance set of M_s/m_b ratios has been included. Figure (IV-14) shows the fusion tree obtained using Ward's method where variables 4 (Pg/Lg_1), 2 (P_2/S_2), and 7 (Rg_1/Lg_1) are used. As was found for the original set of events, the reduced set shows no tendency to cluster into the actual earthquake and explosion populations. Moreover, the clustering process again stabilizes at the four-cluster level, with a large increase in the total error sum of squares required to move to the three-cluster level or

below.

Figures (IV-15) and (IV-16) are the scatterplots showing the cluster distributions obtained at the four-cluster level using Ward's method and procedure RELOCATE. Again, aside from indexing differences, the cluster memberships are identical, which implies an extremely stable cluster grouping. In the same manner described above, the grouping generated by RELOCATE was obtained using a five cluster initial partition.

The fusion tree obtained where the smallest variance set of M_s/m_b ratios together with variables 4 (Pg/Lg_1) and 7 (Lg_1/Rg_1) have been included is given in Figure (IV-17). With M_s/m_b included, the tendency to cluster into the explosion and earthquake populations is much greater. Two earthquake clusters (one containing an explosion) form at the five cluster level. These two clusters fuse at the four-cluster level to form an earthquake cluster with one misplaced explosion (Stoat). However, this earthquake cluster fuses with the larger explosion cluster at the two-cluster level. Two clusters remain at the two-cluster level, where one consists exclusively of explosions and the other cluster contains both earthquakes and explosions. Scatterplots of M_s/m_b versus ratio 4 (Pg/Lg_1) are shown in Figures (IV-18) and (IV-19). Aside from indexing differences, the cluster memberships are identical with one exception. In the cluster grouping generated by RELOCATE, the explosion which is misplaced in the grouping generated

by Ward's method is correctly placed into an explosion cluster.

CHAPTER V

CONCLUSIONS AND DISCUSSION

Discriminant Analyses

The discriminant analysis procedures BMD04M and BMD07M successfully duplicate the results cbtained by Booker and Mitronovas (1964). Also verified in this study are the relative discriminating powers of each of the seven energy ratios when used alone. The ratios, together with their F-statistics are listed in Table (V-1), both for the full set of earthquakes and explosions and for the reduced set where the M_s/m_b ratios have been included. The F-statistics may be used to test the null hypothesis that the two population means are equal. In general, the large values found for the F-statistics support the alternative hypothesis that the two population means are unequal.

Booker and Mitronovas (1964) claim an average overall probability of correct classification of 0.85. In this study, under the minimax assumption that the prior probability of an observation originating in either population is 0.5, the average probability of correct classification is 0.79 where the entire set of events is used, and 0.78 where the reduced set is used. Where the

quadratic procedure is used, the average probability of correct classification is 0.94 for the full set of events and 0.98 for the reduced set of events. Because of the small size of the earthquake sample, however, the result obtained for the reduced set of events by the quadratic procedure should probably not be taken as representative of the true probability of classification.

In addition to studying the discriminant potential of each variable when considered alone, the marginal contributions of the variables have been studied using the stepwise algorithm of BMD07M. In the case of the energy ratios, once (Pg/Lg_1) has been entered at the first step, (P_1/S_1), (Pg/B), and $(Pg/[Lg_1+Rg_1])$ lose much of their initial discriminating power because they are all rather strongly correlated with (Pg/Lg_1). As a result, once (Pg/Lg_1) has been entered, the other three variables provide little additional discriminating information.

The estimated M_s/m_b ratios are not highly correlated with any of the Booker and Mitronovas energy ratios. However, this may not be true for observed M_s/m_b ratios, because the M_s/m_b ratios used in this study are randomly generated within specified variance constraints that characterize several M_s/m_b observations at NTS.

On the other hand, this study shows the enormous power of the M_s/m_b discriminant, even in the pessimistic case where the wide variance of M_s/m_b set 4 is used. Using set 4, the average probability of correct classification is

boosted from 0.78, where only the short-period energy ratios are used, to 0.96. This probability increases to nearly 1.00 for the remaining sets of M_s/m_b ratios.

Another useful result obtained in this study is the successful classifications obtained using the quadratic discriminant analysis procedure. The results obtained using this procedure are found to be at least as good (in terms of the posterior classification probabilities) as those from the linear procedures for the original set of events, and perhaps even better for the reduced set. Some caution must be exercised, however, both in the situation where the sample sizes are small, and where the actual distribution of events differs from the assumption of normality. If a multivariate model is not appropriate for describing the data, misleading classifications might be obtained for observations located at points corresponding to the tails of the distributions where the probability densities are small. On the other hand, where the multivariate normal distribution is a valid model, the quadratic procedure provides the best classification if the covariance matrices differ markedly.

An obvious flaw exists in the selection of events for this study. As shown in Table (IV-1), the explosion data set is drawn primarily from events at NTS, with the [illegible] of Gnome, which was detonated near Carlsbad, New [illegible] however, the earthquake data set is drawn from [illegible] States and Baja California. As

a result, for the underground explosion population, there is a great deal of consistency in the source locations, source depths, and travel paths to stations. The earthquake population, on the other hand, is very poorly constrained in all these respects. This suggests that the discriminant analysis results presented here might be improved a great deal by using a more controlled earthquake sample population. This suggestion is supported by the fact that the Booker and Mitronovas energy ratios and the M_s/m_b discriminant both reflect similar source properties. Energy ratios drawn on a properly constrained set of earthquakes could be expected to discriminate at least as successfully as the M_s/m_b ratios did in this study. Data for a set of properly constrained earthquakes are needed to test this suggestion.

In addition to the discriminants employed in this study, any other discriminant may be readily incorporated into the linear discriminant methods outlined here. The stepwise procedure used in BMD07M makes this approach particularly attractive, in that it may be used to find those sets of variables where each one is a powerful discriminant alone, and where each one is maximally uncorrelated with the other variables in the set. Such a set of nearly independent variables provides optimal classifications because errors introduced in any given variable tend, on the average, to be compensated by the others.

The problem of missing data may be solved quite easily in a discriminant analysis scheme. If several events lack measurements on one or more variables, the missing values may be replaced by the value corresponding to the point of intersection between the hyperplane which optimally separates the two populations and the line connecting the two population means. This point may be found by generating a discriminant analysis where the events containing missing data values have been omitted. In such a procedure, the variables corresponding to the missing values do not contribute to the discriminant analysis for events with missing data. Instead, these events are classified using only the variables for which measurements are available.

Cluster Analyses

From the results obtained above, it may be seen that cluster analysis represents a less useful tool for classification of seismic events than linear and quadratic discriminant analysis. The reasons for this can probably be traced to the fact that, in the case of the discriminant analyses, the group memberships and the number of groups are specified as constraints on the problem. The purpose of the discriminant analysis is to provide the linear or quadratic function which best separates the two populations, given the

input information as to the number of groups and the sample memberships. In the case of cluster analysis, no such prior information is available; the data alone must suggest the appropriate partitioning and the optimum number of groups. In this way, cluster analysis best operates as a hypothesis generator. That is, where no prior information is available about the data, cluster analysis methods might be used to generate information about intrinsic group structure. This information could then serve as the input to a discriminant analysis scheme. Such a strategy is of little use in the present study, however, because the group attributes of the data are already known, and are not those generated by the cluster analyses.

One interesting result which did arise from the cluster analyses generated in this study is the tendency for the data to form stable groups at the four-cluster level. No explanation could be found for this phenomenon. No correlations were observed between the cluster formation and the event source locations, average source-receiver distances, or body wave magnitudes. It may be possible that these cluster formations result from similar source depths for those events in the same cluster, or from some effect which may be traced to the azimuthal distribution of stations. However, no information is available for these events concerning either the stations used to obtain the energy ratios or the earthquake focal depths.

Finally, the incorrect classifications obtained using

cluster analysis in all but the case where the M_s/m_b ratios are included indicates that cluster analysis provides useful groupings only where the between-group scatter is significantly greater than the within-group scatter, or where the groups are highly disjoint.

Suggestions for Improving Discrimination Capabilities based on Recent Advances in Seismic Instumentation and Signal Processing Techniques

Improved discrimination capabilities may be provided by utilizing several recent advances in seismic instrumentation and signal processing techniques. Because new discriminants are readily incorporated into the framework of the discriminant analysis techniques presented in this paper, it is possible to construct training sets of data and to upgrade them at a later time as new discriminants are developed.

Seismic instrumentation has been greatly improved in recent years. A primary example is the development of the Seismic Research Observatory (SRO) network of which 13 stations are now in operation world-wide. The instruments used in the SRO network are characterized by a very broad dynamic range, so that a wide magnitude range of events may be recorded without altering the frequency response

characteristics of the system. Because the SRO seismometers are situated in deep boreholes, exceptionally high signal to noise ratios are realized. The data recorded at the SRO stations are digitized and stored on magnetic tape volumes. This compact means of data storage allows for convenient retrieval of data for subsequent processing. In particular, the on-site conversion of analog data to digital format allows quite sophisticated data processing techniques to be applied routinely and economically.

A seismic network such as the SRO network should prove invaluable in the seismic discrimination problem. One of the reasons that the M_s/m_b discriminant has been so successful is that it may be easily and routinely calculated. The digital format of the SRO data allows more sophisticated discriminants to be calculated as routinely as M_s/m_b, and moreover, because such measurements may be completely automated, the introduction of human error in measurement and digitization may be minimized.

By using seismograms from a system such as the SRO network, it is possible to incorporate discriminants based on frequency domain analyses into the classification procedure. Spectral discriminants have been suggested as a means for improving discrimination capabilities (Weichert, 1971; SIPRI, 1968). By using spectral ratios in the discriminant analysis procedures it should be possible to combine the effects of differences in energy partitioning at the source with those resulting rom different source

dimensions and time histories for the two types of events.

A facility has been provided to calculate these spectral ratios in the seismoprint program described in Appendix C. The seismoprint program generates spectral analyses at two different levels. At the first level, the program obtains a spectrogram for the event being studied. The spectrogram consists of a two-dimensional contour display of power spectral density as a function of time (or velocity) and frequency. At the next level, spectra within selected time (or velocity) windows may be obtained by averaging the spectra used to obtain the event spectrogram. In addition, these spectra may be generated over any frequency range of interest within the available bandwidth. The spectra may then be used in the formation of spectral ratios.

An area which shows considerable promise is the development of a classification scheme based on the event spectrogram. This is especially interesting because spectral information in the entire seismogram could be utilized in a cluster analysis or pattern recognition scheme. Spectrograms are shown in Figures (V-1) and (V-2) for the Shoal explosion and the Fallon earthquake. The spectrograms for these events exhibit remarkable differences in the high frequency spectral content and in the relative partitioning of energy between the body and surface wave radiation fields.

Application of pattern recognition techniques to the

event spectrogram has received little attention in seismology. However, in light of the possibility of automating a pattern recognition scheme using digitized data such as that available from the SRO network, further research in this area is warranted.

REFERENCES

Aki, K., Scaling law of seismic spectrum, J. Geophys. Res., 72, 1217 (1967).

Alexander, S. S., and J. W. Lambert, Long-period Rayleigh wave spectra from earthquakes and explosions, Seismic Data Laboratory Report No. 307, Teledyne Geotech, Alexandria, Va. (1973).

Anderberg, M. R., Cluster Analysis for Applications, Academic Press, New York (1973).

Anderson, T. W., An Introduction to Multivariate Statistical Analysis, J. Wiley and Sons, New York (1958).

Anderson, T. W., and R. R. Badahur, Classification into two multivariate normal distributions with different covariance matrices, Annals of Math. Stat., 33, 420 (1962).

Archambeau, C. B., General theory of elastodynamic source fields, Rev. Geophys., 6, 241 (1968).

Archambeau, C. B., and C. Sammis, Seismic radiation from explosions in prestressed media and the measurement of tectonic stress in the Earth, Rev. Geophys., 8, 473 (1970).

Archambeau, C. B., The theory of stress wave radiation from explosions in prestressed media, Geophys. J. R. Astr. Soc., 29, 329 (1972).

Basham, P. W., Canadian magnitudes of earthquakes and nuclear explosions in south-western North America, Geophys. J. R. Astr. Soc., 17, 1 (1969).

Bath, M., Introduction to Seismology, J. Wiley and Sons, New York (1970).

Bell, A. G. R., and S. S. Alexander, A digital technique for detection of multiple seismic events, Paper presented before the Annual Spring Meeting of the American Geophysical Union, May, 1977, Washington, D. C. (1977).

Ben-Menahem, A., Radiation of seismic surface waves from a finite moving source, Bull. Seism. Soc. Am., 51, 401 (1961).

Ben-Menahem, A., and D. G. Harkrider, Radiation patterns of seismic surface waves from buried dipolar point sources in a flat stratified earth, J. Geophys. Res., 69, 2605 (1964).

Blackman, R. B., and J. W. Tukey, The Measurement of Power Spectra, Dover Books, New York (1958).

BMD: Biomedical Computer Programs, Health Sciences Computing Facility, University of California, Los Angeles (1976).

Bolt, B. A., Nuclear Explosions and Earthquakes: The Parted Veil, W. H. Freeman, San Francisco (1976).

Booker, A., and W. Mitronovas, An application of statistical discrimination to classify seismic events, Bull. Seism. Soc. Am., 54, 961 (1964).

Brune, J. N., Tectonic stress and the spectra of seismic shear waves from earthquakes, J. Geophys. Res., 75, 4997 (1970).

Burridge, R., and L. Knopoff, Body force equivalents for seismic dislocations, Bull. Seism. Soc. Am., 54, 1875 (1964).

Capon, J., R. J. Greenfield, and R. T. Lacoss, Long-period signal processing results for the large aperture seismic array, Geophysics, 34, 305 (1969).

Carpenter, E. W., Teleseismic signals calculated for underground, underwater, and atmospheric explosions, Geophysics, 32, 17 (1967).

Chinnery, M. A., Some physical aspects of earthquake source mechanism, J. Geophys. Res., 65, 3852 (1960).

Chinnery, M. A., The deformation of the ground around surface faults, Bull. Seism. Soc. Am., 51, 355 (1961).

Cooley, J. W., P. A. W. Lewis, and P. D. Welch, The Fast Fourier Transform and its applications, IBM Research Paper RC-1743, IBM Watson Research Center, Yorktown Heights, New York (1967).

Douglas, A., J. A. Hudson, and V. K. Kembhavi, The relative excitation of seismic surface and body waves by point sources, Geophys. J. R. Astr. Soc., 23, 451 (1971).

Everitt, B., Cluster Analysis, Heineman Educational Books, London (1974).

Evernden, J. F., Magnitude determinations at regional and near-regional distances in the United States, Bull. Seism. Soc. Am., 57, 591 (1969).

Fisher, R. A., The use of multiple measurements in taxonomic problems, Ann. Eugen., 9, 238 (1936).

Gilbert, F., Excitation of the normal modes of the Earth by earthquake sources, Geophys. J. R. Astr. Soc., 22, 223 (1970).

Gilbert, F., The relative efficiency of earthquakes and explosions in exciting body waves and surface waves, Geophys. J. R. Astr. Soc., 33, 487 (1973).

Hanks, T. C., and M. Wyss, The use of body wave spectra in the determination of seismic source parameters, Bull. Seism. Soc. Am., 62, 561 (1972).

Harkrider, D. G., Surface waves in multilayered elastic media: I. Rayleigh and Love waves from buried sources in a multilayered elastic half-space, Bull. Seism. Soc. Am., 54, 627 (1964).

Hasegawa, H. S., Analysis of amplitude spectra of P-waves from earthquakes and underground explosions, J. Geophys. Res., 77, 3081 (1972).

Hodgson, J. H., and A. E. Stevens, Seismicity and earthquake mechanisms, in Research in Geophysics, Hugh Odishaw, ed., v. 2, MIT Press, Cambridge Massachussetts (1964).

Honda, H., The generation of seismic waves, in A Symposium on Earthquake Mechanism, J. H. Hodgson, ed., Pub. Dom. Obs. Ottowa, 24, 329 (1961).

Kanasewich, E. R., Time Sequence Analysis in Geophysics, University of Alberta Press, Edmonton, Alberta (1975).

Keilis-Borok, V. I., Investigation of the mechanism of earthquakes, Soviet Res. in Geophys., 4, American Geophysical Union, Consultant's Bureau, New York (1960).

Knopoff, L., and F. Gilbert, First motion from seismic sources, Bull. Seism. Soc. Am., 50, 117 (1960).

Koopmans, L. H., The Spectral Analysis of Time Series, Academic Press, New York (1974).

Lapp, R., The Voyage of the Lucky Dragon, Harper, New York (1958).

Liebermann, R. C., and P. W. Pomeroy, Relative excitation of surface waves by earthquakes and underground explosions, J. Geophys. Res., 74, 1575 (1969).

Marshall, P. D., Aspects of the spectral differences between earthquakes and underground explosions, Geophys. J. R. Astr. Soc., 20, 397 (1970).

Marshall, P. D., and P. W. Basham, Discrimination between earthquakes and nuclear explosions using an improved M_s scale, Geophys J. R. Astr. Soc., 28, 431 (1972).

Mitronovas, W., Identification of nuclear explosions using ratios of integrated seismic phases, Research Memorandum (DATDC Report No. 84), United Electrodynamics, Inc., Alexandria Va. (1963).

Press, F., A. Ben-Menahem, and M. N. Toksoz, Experimental determination of earthquake fault length and rupture velocity, J. Geophys. Res., 66, 3471 (1961).

Press, F., G. Dewart, and R. Gilman, A study of diagnostic techniques for identifying earthquakes, J. Geophys. Res., 68, 2909 (1963).

Rao, C. R., Linear Statistical Inference and its Applications, 2nd ed., J. Wiley and Sons, New York (1973).

Reid, H. F., Physics of the Earth, VI, 100 (1933).

Sharpe, J. A., The production of elastic waves by explosion pressures: I. Theory and empirical field results, Geophysics, 7, 155 (1946).

SIPRI, Seismic methods for monitoring underground explosions, International Peace Research Institute, Stockholm (1968).

Stacey, F. D., Physics of the Earth, J. Wiley and Sons, New York (1969).

Stauder, W. S., The focal mechanism of earthquakes, in Advances in Geophysics, Academic Press, New York, 9, 1 (1962).

Steketee, I. A., Some geophysical applications of elasticity theory of dislocations, Can. J. of Phys., 36, 1168 (1958).

Stevens, A. E., World-wide earthquake mechanism, in The Earth's Crust and Upper Mantle, Geophysical Monograph 13, American Geophysical Union, Washington, D. C. (1969).

Thirlaway, H. I. S., Diagnosing underground explosions and earthquakes, Contemp. Phys., 9, 17 (1968).

Tjostheim, D., Autoregressive representation of seismic P-wave signals with application to the problem of short-period discriminants, Geophys. J. R. Astr. Soc., 43, 269 (1975).

Turnbull, L. S., Determination of seismic source parameters using far-field surface wave spectra, Ph.D. Thesis, The Pennsylvania State University (1976).

von Seggern, D. H., Joint magnitude determination and analysis of variance for explosion magnitude estimates, Seismic Data Laboratory Report No. 286, Teledyne Geotech, Alexandria Va. (1972).

Weichert, D. H., Short-period spectral discriminant for earthquake-explosion differentiation, Zeits. fur Geophysik, 37, 147 (1971).

Whalen, A. D., Detection of Signals in Noise, Academic Press, New York (1971).

Wishart, D., Clustan 1C User Manual, The Clustan Project, University College, London (1975).

Wyss, M., T. C. Hanks, and R. C. Liebermann, Comparison of P-wave spectra of underground explosions and earthquakes, J. Geophys. Res., 76, 2715 (1971).

APPENDIX A

THREE CLASSIFICATION CRITERIA OFTEN USED IN DISCRIMINANT ANALYSES

Decision Probabilities and Cost Functions

For all of the criteria which follow, if the random vector $\vec{x}$ falls within some region R_i (determined from the criterion), it is identified with the population A_i. It is possible to formulate various decision probabilities in terms of integrals of the appropriate likelihood functions over the regions R_i. Let D_{ij} represent the decision that the event in question originates in the i^{th} population when it actually comes from the j^{th} population. In terms of the hypotheses H_i, this is equivalent to choosing H_i when H_j is actually true. The probability of making this decision is given by:

$$Pr\{D_{ij}\} = \int_{R_i} p_j(\vec{x})\, dV. \qquad (A - 1)$$

This is illustrated for the error probabilities $Pr\{D_{12}\}$ and $Pr\{D_{21}\}$ for the two group univariate case in figure (A-1). The average probability of correctly classifying the random vector $\vec{x}$ is given by:

$$P_c = \sum_{i=1}^{2} P_r\{D_{ii}\} P_r\{H_i\}. \tag{A - 2}$$

Similarly, the average probability of misclassifying x is given by

$$P_e = \sum_{i=1}^{2} \sum_{j=1}^{2} P_r\{D_{ij}\} P_r\{H_i\} = 1 - P_c. \tag{A - 3}$$

In many cases it is desirable to weight, or assign costs to the various decisions. If C_{ij} denotes the cost associated with the decision D_{ij}, the average cost (i.e., the cost function) associated with classifying the random vector $\vec{x}$ is given by:

$$\bar{C} = \sum_{i=1}^{2} \sum_{j=1}^{2} C_{ij} P_r\{D_{ij}\} P_r\{H_j\}. \tag{A - 4}$$

In general, the cost associated with making a correct decision is set equal to zero. This convention will be followed here.

The Bayes Criterion

The Bayes criterion dictates that $\vec{x}$ be classified in such a way that the cost function is minimized, which is equivalent to determining the regions R_i so that $\bar{C}$ is minimized. For the two population case, assuming that C_{11} =

$C_{22} = 0,$

$$\bar{C} = C_{12}\Pr\{D_{12}\}\Pr\{H_2\} + C_{21}\Pr\{D_{21}\}\Pr\{H_1\} \tag{A - 5}$$

Letting $\Pr\{D_{21}\} = 1 - \Pr\{D_{11}\}$, and $\Pr\{H_2\} = 1 - \Pr\{H_1\}$, the average cost may be rewritten:

$$\bar{C} = C_{21}\Pr\{H_1\} + \int_{R_1}\left[C_{12}(1-\Pr\{H_1\})\,p_2(\vec{x}_o) - C_{21}\Pr\{H_1\}\,p_1(\vec{x}_o)\right]dV. \tag{A - 6}$$

Because R_1 is the only variable quantity in (A-6), in order to minimize C, R_1 must be chosen to include only those values $\vec{x}_o$ for which

$$C_{12}(1-\Pr\{H_1\})\,p_2(\vec{x}_o) - C_{21}\Pr\{H_1\}\,p_1(\vec{x}_o) \le 0. \tag{A - 7}$$

This result may be stated in terms of the likelihood ratio test for H_1 as:

Choose H_1 if:

$$p_1(\vec{x}_o)/p_2(\vec{x}_o) > C_{12}(1-\Pr\{H_1\})/C_{21}\Pr\{H_1\},$$

choose H_2 otherwise.

It should be noted that if $C_{12} = C_{21}$ and $C_{11} = C_{22} = 0$, the average cost reduces to the average probability of error. This special case of the Bayes criterion which minimizes P_e

is known as the minimum error, or maximum a posteriori probability criterion.

The Neymann-Pearson Criterion

The Neymann-Pearson criterion is quite useful in the two population case, in that it does not require either costs or prior probabilities. Under this criterion, the regions R_i are chosen so that $Pr\{D_{ii}\}$ is maximized, while $Pr\{D_{ij}\}$ is set to a specified threshold value, say z. Since $Pr\{D_{ii}\} = 1 - Pr\{D_{ji}\}$, this procedure is equivalent to minimizing $Pr\{D_{ji}\}$. Suppose $Pr\{D_{12}\}$ is set at z. Then $Pr\{D_{21}\}$ is to be minimized subject to the constraint: $Pr\{D_{12}\} = z$. Introducing the Lagrange multiplier λ, the quantity to be minimized is given by Q, where:

$$Q = Pr\{D_{21}\} + \lambda\, Pr\{D_{12}\} \qquad (A-8)$$

Expression (A-8) is the Bayes cost where $C_{11} = C_{22} = 0$, $C_{12}(1-Pr\{H_1\}) = \lambda$, and $C_{21}Pr\{H_1\} = 1$, so that R_1 is chosen such that:

$$\lambda\, p_2(\bar{x}) - p_1(\bar{x}) \leq 0 \qquad (A-9)$$

or in terms of the likelihood ratio test for H_1:

Choose H_1 if:

$$p_1(\vec{x}_o)/p_2(\vec{x}_o) \geq \lambda,$$

choose H_2 otherwise.

The Minimax Criterion

Another criterion which is widely used where the prior probabilities are unknown is the minimax criterion, which dictates that the decision regions be chosen in such a way as to minimize the maximum average cost. That is, the minimax criterion implicitly specifies that the prior probabilities be chosen to correspond to the maximum on the Bayes cost curve. This is illustrated in figure (A-2).

If the Bayes cost is viewed as a function of $Pr\{H_i\}$ with the C_{ij} and $Pr\{D_{ij}\}$ fixed, it has a shape similar to that shown in figure (A-2). The Neymann-Pearson criterion has the effect of linearizing the average cost. The Neymann-Pearson cost is quite close to the Bayes cost for values of $Pr\{H_i\}$ close to A. However, as $Pr\{H_i\}$ approaches and exceeds B, the Neymann-Pearson cost becomes quite large. An alternative is to choose the decision regions R_1 and R_2 so the maximum possible cost is minimized. This is done by the minimax criterion.

The maximum on the Bayes cost curve is found by

differentiating the Bayes cost with respect to the prior probability $Pr\{H_1\}$, which is considered as an independent variable, β. The result is then set to zero:

$$\frac{\partial \bar{C}}{\partial \beta} = \frac{\partial}{\partial \beta}\left[C_{12}\, P_r\{D_{21}\}(1-\beta) + C_{21}\, P_r\{D_{21}\}\beta\right]$$

$$= C_{21}\, P_r\{D_{21}\} - C_{12}\, P_r\{D_{12}\} \qquad (A - 10)$$

$$= 0.$$

The condition for the minimax criterion, then, requires that the decision regions be chosen so that

$$C_{12}\, P_r\{D_{12}\} = C_{21}\, P_r\{D_{21}\} \qquad (A - 11)$$

The minimax cost is found by substituting the result from (A-11) into the expression for the Bayes cost given in (A-6), so that the minimax cost becomes:

$$\bar{C} = C_{21}\, P_r\{D_{21}\} + C_{12}\, P_r\{D_{12}\} \qquad (A - 12)$$

Because $\bar{C}$ determines $Pr\{D_{12}\}$ and $Pr\{D_{21}\}$, the decision regions R_1 and R_2 are those regions satisfying:

$$C_{12}\int_{R_1} p_2(\vec{x})\, dV = C_{21}\int_{R_2} p_1(\vec{x})\, dV. \qquad (A - 13)$$

Again, this result may be expressed in terms of the likelihood ratio test for H_1:

Choose H_1 if:

$$p_1(\vec{x}_o)/p_2(\vec{x}_2) > \lambda_M,$$

choose H_2 otherwise,

where λ_M is chosen to satisfy the minimax criterion. The minimax cost is independent of $Pr\{H_1\}$, and is tangent to the maximum on the Bayes cost curve as shown in figure (A-2).

The minimax cost represents the best criterion where no information is available about the prior probabilities $Pr\{H_i\}$. If $Pr\{H_1\}$ is known to be in some region such as $0 < Pr\{H_1\} < B$, as shown in Figure (A-2), then the Neymann-Pearson criterion may yield a smaller average cost, and be better suited to the problem. However, if the prior probabilities are known, the Bayes criterion always yields the minimum costs.

Each of the criteria discussed above leads to a simple likelihood ratio test in the two population case. Multiple populations may be as easily dealt with by considering the Bayes cost given in (A-6):

$$\bar{C} = \sum_{i=1}^{J} \sum_{j=1}^{J} C_{ij} \, Pr\{D_{ij}\} \, Pr\{H_j\}. \qquad \text{(A-14)}$$

The Bayes criterion requires that this average cost be

minimized, which implies that for any realization of $\vec{x}$, say $\vec{x} = \vec{x}_o$, the hypothesis chosen must be the one which yields the lowest average cost. The cost associated with H_j is given by:

$$C_j = \sum_{i=1}^{g} C_{ji} \Pr\{H_i / \vec{x} = \vec{x}_o\} \tag{A - 15}$$

$\Pr\{H_i | \vec{x} = \vec{x}_o\}$ is given by Bayes' rule:

$$\Pr\{H_i / \vec{x} = \vec{x}_o\} = \frac{p_i(\vec{x}_o) \Pr\{H_i\}}{p(\vec{x}_o)} \tag{A - 16}$$

so that

$$C_j = \sum_{i=1}^{g} C_{ji} \frac{p_i(\vec{x}_o) \Pr\{H_i\}}{p(\vec{x}_o)} \tag{A - 17}$$

Since the marginal density of $\vec{x}$, $p(\vec{x})$, is common to all terms, the decision rule is to choose the hypothesis which minimizes:

$$\lambda_j = \sum_{i=1}^{g} C_{ji} \, p_i(x_o) \Pr\{H_i\} \tag{A - 18}$$

The expression above is the one derived by Whalen (1971). An equivalent expression, the discriminant score, is given by Rao (1973). Given that $\vec{x} = \vec{x}_o$, the discriminant

score for the jth population is given by:

$$S_j = - \sum_{i=1}^{g} C_{ji} \, p_i(\vec{x}_0) \, P_r\{H_i\} \qquad (A - 19)$$

In this case the rule is to choose the hypothesis which maximizes S_j.

APPENDIX B

THE QUADRATIC DISCRIMINANT ANALYSIS PROGRAM, QDISCRIM

Program QDISCRIM may be used to generate a quadratic discriminant analysis for a set of sample vectors from as many as ten groups. The maximum number of variables which may be included in the analysis is also limited to ten. The maximum number of sample vectors within each group is limited to 100. QDISCRIM classifies the input observations on the basis of the relative magnitudes of the equivalent discriminant score (Rao, 1973; Whalen, 1971) for each group. Each sample vector is classified into that group for which its equivalent quadratic discriminant score is greatest. The sample vectors are assumed to have a multivariate normal distribution within each group.

The quadratic procedure has the advantage that no assumption need be made concerning the equality of the covariance matrices for each group. This assumption is implicit in all linear discriminant analysis techniques. On the other hand, the quadratic procedure relies most heavily on the assumption of normality because the boundary for the classification regions often coincides with areas far from the means of the respective likelihood functions, where the

probability densities are small (Anderson and Badahur, 1962). Because the sample covariance matrices are calculated for each group individually, caution must be exercised to make sure that the number of observation vectors in each group exceeds the number of components in the observation vector so that the sample covariance matrices are non-singular.

Computational Procedure

Let $\vec{x}^1{}_1,\ldots,\vec{x}^1{}_{N1},\ldots,\vec{x}^g{}_1,\ldots,\vec{x}^g{}_{Ng}$ represent observation vectors from g groups, where N_i is the number of observations in the i^{th} group. The first step in QDISCRIM consists of estimating the mean vectors ($\vec{\bar{x}}_i$) and the sample covariance matrices (S_i) for each group. For the i^{th} group, $\vec{\bar{x}}_i$ and S_i are given by:

$$\vec{\bar{x}}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} \vec{x}_j^{\,i},$$

$$S_i = \frac{1}{N_i-1}\sum_{j=1}^{N_i} (\vec{x}_j^{\,i} - \vec{\bar{x}}_i)(\vec{x}_j^{\,i} - \vec{\bar{x}}_i)' \qquad (B - 1)$$

Because the sample covariance matrices are positive definite, the determinant of S_i is given by

$$|S_i| = \prod_{j=1}^{p} \lambda_j^i . \qquad (B - 2)$$

where λ^{i}_{j} is the j^{th} eigenvalue of the sample covariance matrix for the i^{th} group. The eigenvalues are obtained using the Jacobi method employed in the PSU library subroutine DMXDJA. Finally, each sample covariance matrix is inverted using the PSU library subroutine DMXINV.

The next step in the QDISCRIM procedure involves evaluating the equivalent quadratic discriminant score for each observation vector, and classifying the vector into that group for which the quadratic score is greatest. This is equivalent to classifying the observation into the group which maximizes the probability that the hypothesis H_i (i.e., that $\vec{x}_o$ originates in the j^{th} group) is true. This probability may be formulated in terms of the likelihood functions $p_i(\vec{x})$ by applying Bayes rule:

$$Pr\{H_i / \vec{x} = \vec{x}_o\} = \frac{p_i(\vec{x}_o)\, Pr\{H_i\}}{p(\vec{x}_o)} \qquad (B - 3)$$

In (B-3), $p_i(\vec{x}_o)$ is the likelihood function (the conditional probability density function of $\vec{x}$, given that H_i is true) for the i^{th} population evaluated at $\vec{x} = \vec{x}_o$, where $\vec{x}_o$ is a specific realization of the random vector $\vec{x}$ corresponding to each of the sample vectors. $p(\vec{x}_o)$ is the marginal probability density of $\vec{x}$, also evaluated at $\vec{x}_o$. $Pr\{H_i\}$ is the prior probability that $\vec{x}_o$ originates in the i^{th} group, and, in the case of QDISCRIM, may be specified by the user. The likelihood functions $p_i(\vec{x})$ are estimated by:

$$p_1(\vec{x}) = \frac{1}{(2\pi)^{p/2}|S_1|^{1/2}} \exp\left\{-\tfrac{1}{2}(\vec{x}-\bar{x}_1)'S_1^{-1}(\vec{x}-\bar{x}_1)\right\}$$

$$p_2(\vec{x}) = \frac{1}{(2\pi)^{p/2}|S_2|^{1/2}} \exp\left\{-\tfrac{1}{2}(\vec{x}-\bar{x}_2)'S_2^{-1}(\vec{x}-\bar{x}_2)\right\} \tag{B - 4}$$

and the marginal density of $\vec{x}$ is given by:

$$p(\vec{x}) = \sum_{i=1}^{g} p_i(\vec{x})\, P_A\{H_i\} \tag{B - 5}$$

The classification strategy used in QDISCRIM involves identifying $\vec{x}_0$ with the group which maximizes the right side of (B-3). Since $p(\vec{x})$ and the factor involving 2 are common to all groups, they may be deleted from the comparison. Also, because the natural logarithm is a strictly increasing function, it is sufficient to classify $\vec{x}_0$ into the group which maximizes:

$$\mathcal{S} = \log(P_A\{H_i\}) - \tfrac{1}{2}(\vec{x}-\bar{x}_i)'S_i^{-1}(\vec{x}-\bar{x}_i) - \tfrac{1}{2}\log|S_i|. \tag{B - 6}$$

The expression given in (B-6) is the equivalent discriminant score (Rao, 1973; Whalen, 1971). For each sample vector, QDISCRIM evaluates (B-6) for each group, then classifies the sample vector into the group for which $\mathcal{S}$ is maximal. Finally, QDISCRIM evaluates and lists the classification probabilities given in (B-3) for each observation vector in

each group.

APPENDIX C

THE SEISMOPRINT PROGRAM

The seismoprint program is a comprehensive spectral analysis program developed to analyze seismic signatures in several ways. In the context of the earthquake/underground explosion discrimination problem, it may be used to generate energy ratios, either from spectra over the entire bandwidth of the recording instrument, or from spectra within selected frequency bands. In addition, the seismoprint program generates the event spectrogram, the seismological analogue to the voiceprint used in linguistics analysis. This suggests that pattern recognition techniques based on the event spectrogram might also be adapted to the discrimination problem, a possibility that remains largely unexplored.

Program Description

The seismoprint program consists of a main driver and three primary subroutines. The main driver reads the input and control parameter cards, and directs control to each of

the three primary subroutines depending on the tasks the program is to accomplish.

Subroutine SPXFRM reads the input seismogram from tape and generates the power spectra of consecutive, possibly overlapping segments of the seismogram at uniform time increments. The power spectrum at each time point is obtained by removing the mean and linear trend from the desired segment of data, applying a cosine taper to the ends of the window, then extending the window with zeros to the specified power of two. The finite Fourier transform of the data is calculated using the Cooley-Tukey fast Fourier transform algorithm, then the periodogram power spectral density estimate is generated by obtaining the squared modulus of the finite Fourier transform. The periodogram estimate may be smoothed at the user's option. These power spectra are then transferred to disk to await subsequent processing.

Subroutine SPCPLT can be called to generate the spectrogram plot in one of two possible ways. The first option provides a two-dimensional printer contour plot where representation of the contour intervals is achieved by overstriking selected print characters, so that the contour levels are indicated by varying ink densities. The second option provides a hidden-line perspective plot which is generated by the PSU library subroutine HIDE. In both cases, the user has the option of normalizing the power spectrum at each time index with respect to the maximum in

that spectrum. This type of display is quite useful for studying multiple events, in that spectral minima tend to persist throughout the spectrogram, whereas other effects are randomized by the normalization process (Bell, 1977).

The spectrogram may be displayed either as a function of time and frequency or as a function of velocity and frequency. Where the velocity display is chosen, the event spectra of the time indices closest to the specified velocities are plotted at each velocity index. Such a display introduces a scaling effect so that spectrograms for events with different source-receiver distances may be directly compared.

Subroutine SPSPLT generates, at the user's option, plots of both the input seismogram and power spectra of selected time and/or velocity windows. The power spectra for these selected windows are obtained by averaging the spectra of the uniformly spaced segments of the seismogram that are generated by SPXFRM. These spectra may also be saved on disk or tape for subsequent processing, and may be used to generate spectral ratios for a seismic classification scheme.

Properties of Spectral Estimates Obtained by Spectral Averaging

The technique used in subroutine SPSPLT to obtain spectra for the specified time and velocity windows involves averaging the spectra of shorter, possibly overlapping windows that are generated by SPXFRM. This is one procedure suggested by Cooley, Lewis, and Welch (1967) to obtain a consistent power spectral density estimator.

The statistical properties of both the periodogram and weighted covariance power spectral density estimators are considered in some detail by Blackman and Tukey (1959), Cooley, Lewis, and Welch (1967), Koopmans (1974), and Kanasewich (1975). Koopmans (1974) provides a detailed discussion of the distribution of the periodogram spectral estimator under white noise assumptions, and the asymptotic distribution where the input sequence is correlated and non-Gaussian. Koopmans (1974) and Cooley, et al. (1967) demonstrate that, under fairly weak conditions on the length of the data and the smoothness of the true power spectrum, the asymptotic distribution of the periodogram may be used as a reasonable approximation to the true distribution. Under these assumptions, the periodogram power spectral density estimates are approximately independently and identically distributed proportional to chi-square with two degrees of freedom, except at the frequency points

corresponding to $\omega = 0, \pi$, where the estimates are distributed approximately proportional to chi-square with one degree of freedom.

Although bias may be introduced through spectral leakage, which results from either (or both) truncation effects or a rapidly varying true power spectrum, the periodogram estimate is asymptotically unbiased. However, the mean-squared error remains constant in the asymptotic limit (a consequence of the asymptotic chi-square distribution), so that the simple periodogram is not a consistent estimator for the power spectral density. Two techniques are available for generating consistent spectral estimators based on the periodogram. The smoothed periodogram is discussed in detail by Koopmans (1974). The second method, that of spectral averaging, is due to Cooley et al. (1967).

This technique for obtaining a consistent spectral estimate is illustrated in Figure (C-1). The original M sample record is broken into K sections, where each section is L samples long, and is offset from the previous record section by D samples. Periodogram estimates are calculated for each section, and the spectral estimate is obtained by averaging over these periodograms.

The expression for the variance of this estimator involves the lag windows used to weight and truncate the series, so that the weights w(j) must be explicitly included in the following derivations. In the case where the series

is simply truncated, these weights correspond to the simple boxcar window.

The modified periodogram for the pth record segment is given by:

$$I_p(\omega_n) = \frac{L}{2\pi U} \left| A_p(n) \right|^2 \tag{C - 1}$$

where

$$U = \frac{1}{L} \sum_{j=0}^{L-1} W^2(j) \tag{C - 2}$$

and $A_p(n)$ is the finite Fourier transform of the pth windowed record segment, given by:

$$A_p(n) = \frac{1}{L} \sum_{j=0}^{L-1} W(j) X(j) e^{-i\omega_n j} \tag{C - 3}$$

The desired spectral estimate is given by the average over these periodograms:

$$\hat{P}(\omega_n) = \frac{1}{K} \sum_{p=0}^{K-1} I_p(\omega_n). \tag{C - 4}$$

As before, the $I_p(\omega_n)$ are independently and identically distributed proportional to chi-square with two degrees of freedom, except at $\omega_n = 0$ and $\omega_n = \pi$. At these points the $I_p(\omega_n)$ are asymptotically distributed proportional to chi-square with one degree of freedom. Writing $I_p(\omega_n)$ in

terms of the cyclic lagged product of the w(j)x(j) series:

$$I_p(w_n) = \frac{1}{2\pi}\sum_{k=0}^{L-1}\left\{\frac{1}{LU}\sum_{j=0}^{L-1} w(j)\,w(j+k)\,x_p(j)\,x_p(j+k)\right\}e^{-iw_n k}$$

$$= \frac{1}{2\pi}\sum_{k=0}^{L-1} R_p(k)\,e^{-iw_n k} \qquad \text{(C - 5)}$$

the expected value of $\hat{p}(w_n)$ is given by:

$$E\{\hat{p}(w_n)\} = \frac{1}{K}\sum_{p=0}^{K-1} E\{I_p(w_n)\}$$

$$= \int_{-\pi}^{\pi} h_m(w_n - y)\,p(y)\,dy \qquad \text{(C - 6)}$$

where

$$h_m(w) = \frac{1}{2\pi LU}\left|\sum_{j=0}^{L-1} w(j)\,e^{-iw_n j}\right|^2 \qquad \text{(C - 7)}$$

If the true power spectrum $p(w)$ is effectively constant over the bandwidth of $h_m(w)$, $E\{\hat{p}(w_n)\}$ becomes:

$$E\{\hat{p}(w_n)\} \cong p(w_n)\int_{-\pi}^{\pi} h_m(w)\,dw$$

$$\cong p(w_n), \quad \text{since} \quad \int_{-\pi}^{\pi} h_m(w)\,dw = 1. \qquad \text{(C - 8)}$$

which shows that $\hat{p}(w_n)$ is approximately unbiased for sufficiently large L.

The variance of $\hat{p}(w_n)$ is quite difficult to obtain for small L. However, for sufficiently large L (and M), the asymptotic conditions given above are valid to a good degree

of approximation (Koopmans, 1974). The variance of $\hat{p}(\varpi_n)$ is given by:

$$var\{\hat{p}(\varpi_n)\} = var\left\{\frac{1}{k}\sum_{p=0}^{k-1} I_p(\varpi_n)\right\}$$

$$= \frac{1}{k^2}\sum_{p=0}^{k-1}\sum_{q=0}^{k-1} cov\{I_p(\varpi_n)\, I_q(\varpi_n)\}. \qquad (C - 9)$$

Substituting the expression for $I_p(\varpi_n)$ given in (C-4) and rearranging:

$$cov\{I_p(\varpi_n)\, I_q(\varpi_n)\} = E\{I_p(\varpi_n)\, I_{p+j}(\varpi_n)\} - p^2(\varpi_n)$$

$$E\{I_p(\varpi_n)\, I_{p+j}(\varpi_n)\} = \qquad (C - 10)$$

$$= \frac{1}{4\pi^2 L^2 U^2}\sum_{u=0}^{L-1}\sum_{v=0}^{L-1}\sum_{s=0}^{L-1}\sum_{t=0}^{L-1} w(s)\, w(s+u)\, w(t)\, w(t+v) \;\times$$

$$\times\; E\{x_p(s)\, x_p(s+u)\, x_{p+j}(t)\, x_{p+j}(t+v)\}\, e^{-i\varpi_n(u-v)}$$

For sufficiently large L, the series $x_p(j)$ may be viewed as approximately Gaussian and independent. This corresponds to $p(\varpi)$ being approximately constant over $h_m(\varpi)$, and $I_p(\varpi_n)$ being asymptotically distributed proportional to chi-square. Under these conditions, the fourth moment in (C-10) may be expanded. First, since $x_{p+j}(k) = x_p(k+jD)$:

$$E\{X_p(s)X_p(s+u)X_p(t+jD)X_p(t+jD+v)\}$$

$$= E\{X_p(s)X_p(s+u)\}E\{X_p(t+jD)X_p(t+jD+v)\}$$

$$+ E\{X_p(s)X_p(t+jD)\}E\{X_p(s+u)X_p(t+jD+v)\}$$

$$+ E\{X_p(s)X_p(t+jD+v)\}E\{X_p(s+u)X_p(t+jD)\} \quad (C - 11)$$

$$= \sigma^4(\delta_{s,s+u}\,\delta_{t+jD,t+jD+v} + \delta_{s,t+jD}\,\delta_{s+u,t+jD+v}$$

$$+ \delta_{s,t+jD+v}\,\delta_{s+u,t+jD})$$

Substitution of the delta-functions in (C-11) into (C-10) reduces the number of summations from four to two. After carrying through the algebra, (C-10) reduces to:

$$Cov\{I_p(w_n)\,I_{p+j}(w_n)\} = \frac{\sigma^4}{4\pi^2 L^2 U^2}\left\{\sum_{k=0}^{L-1} W(k)W(k+jD)\right\}^2$$

$$= \frac{P^2(w_n)}{L^2U^2}\left\{\sum_{k=0}^{L-1} W(k)\,W(k+jD)\right\}^2 \quad (C - 12)$$

Three points are worth noting here. In the case where $D \geq L$, $w(k+jD) = 0$ for $k=0,\ldots,L-1$, so that $I_p(w_n)$ and $I_{p+j}(w_n)$ are uncorrelated. This situation occurs where the L-sample sections do not overlap. The second point concerns the variance of the $I_p(w_n)$, which may be found from (C-12) by setting j equal to zero:

$$var\{I_p(w_n)\} = \frac{\sigma^4}{4\pi^2L^2U^2}\left\{\sum_{k=0}^{L-1} W^2(k)\right\}^2$$

$$= P^2(w_n). \quad (C - 13)$$

This is indeed the result found above for the simple periodogram. Finally the correlation between $I_p(w_n)$ and $I_{p+j}(w_n)$ may be expressed in terms of the w(k) in (C-12) and (C-13). Since $\mathrm{corr}\{I_p(w_n)I_{p+j}(w_n)\}$ is given by:

$$\rho(j) = \mathrm{corr}\left\{I_p(w_n)\, I_{p+j}(w_n)\right\}$$

$$= \frac{\mathrm{cov}\left\{I_p(w_n)\, I_{p+j}(w_n)\right\}}{\mathrm{var}\left\{I_p(w_n)\right\}} \qquad \text{(C - 14)}$$

$\rho(j)$ may be written:

$$\rho(j) = \frac{\left[\sum_{k=0}^{L-1} W(k)\,W(k+jD)\right]^2}{\left[\sum_{k=0}^{L-1} W^2(k)\right]^2} \qquad \text{(C - 15)}$$

$\rho(j)$ may now be substituted into (C-9), so that $\mathrm{var}\{\hat{p}(w_n)\}$ becomes

$$\mathrm{var}\left\{\hat{p}(w_n)\right\} =$$

$$= \frac{\mathrm{var}\left\{I_p(w_n)\right\}}{K}\left\{1 + 2\sum_{j=1}^{K-1}\left(1 - j/K\right)\rho(j)\right\} \qquad \text{(C - 16)}$$

where $D \geq L$ (i.e., the sections do not overlap), the variance of $\hat{p}(w_n)$ is simply $(1/K)\mathrm{var}\{I_p(w_n)\} = p^2(w_n)/K$. Because this variance goes to zero as K becomes large, $\hat{p}(w_n)$ is a consistent estimator. If M becomes large and L is held constant, then K also becomes large, with the result that:

$$\lim_{M \to \infty} var\left\{\hat{P}(w_n)\right\} = \lim_{K \to \infty} \frac{P^2(w_n)}{K} = 0 \qquad (C - 17)$$

Where D < L, the series involving $\rho(j)$ must be included. However, because the $\rho(j)$ decrease quickly with j (Cooley, et al., 1967), the consistent property of the estimator is retained, and in some cases enhanced. For fixed M and L, the presence of the $\rho(j)$ in the numerator of (C-16) tends to inflate the variance. On the other hand, because the segments overlap, K is larger here than it is in the case of non-overlapping sections, which results in a net decrease in the variance. Finally, it should be noted that for $w_n = 0$ or $w_n = \pi$, the usual doubling in the variance occurs due to the loss of a degree of freedom in the approximating chi-squared distribution.

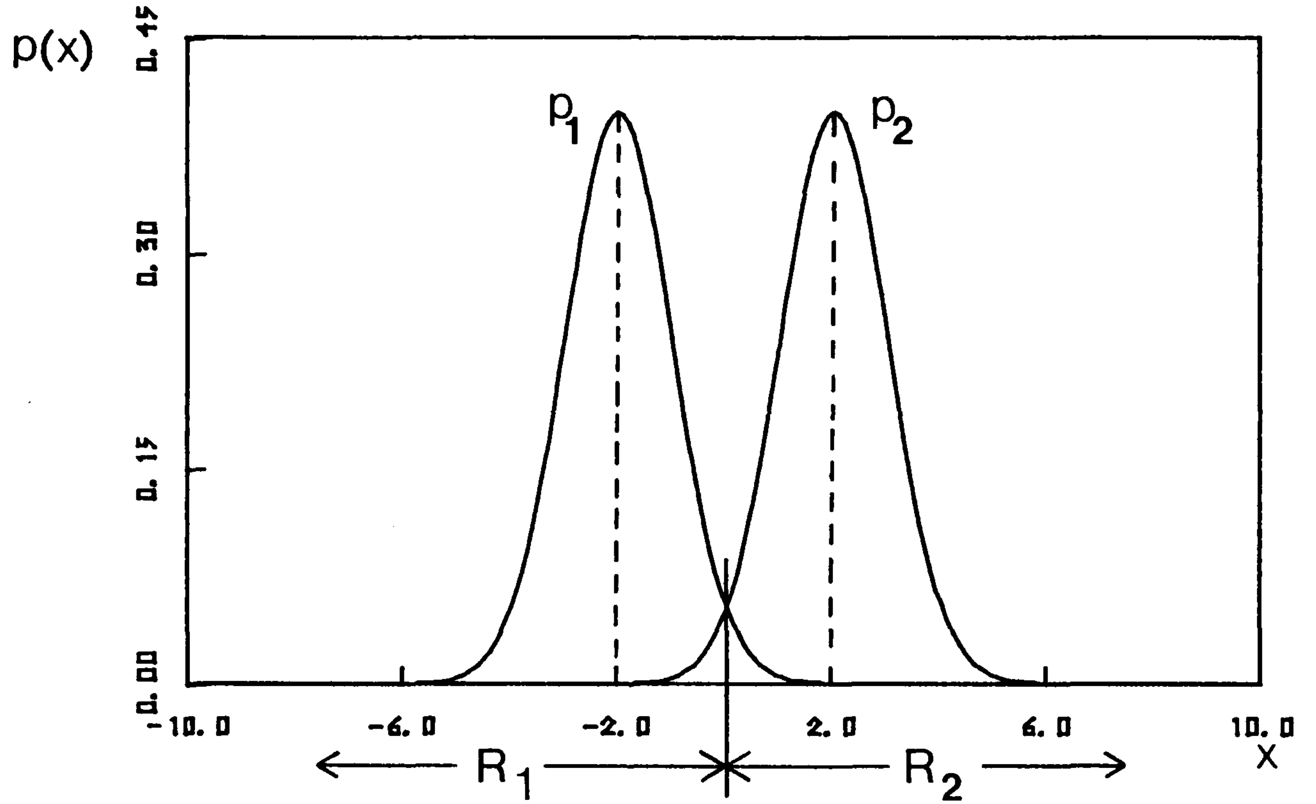


FIGURE (III - 1) Schematic representation of the classification regions in the two-group univariate linear discriminant analysis problem. The likelihood functions for the two groups are denoted by p_1 and p_2, the classification regions by R_1 and R_2.

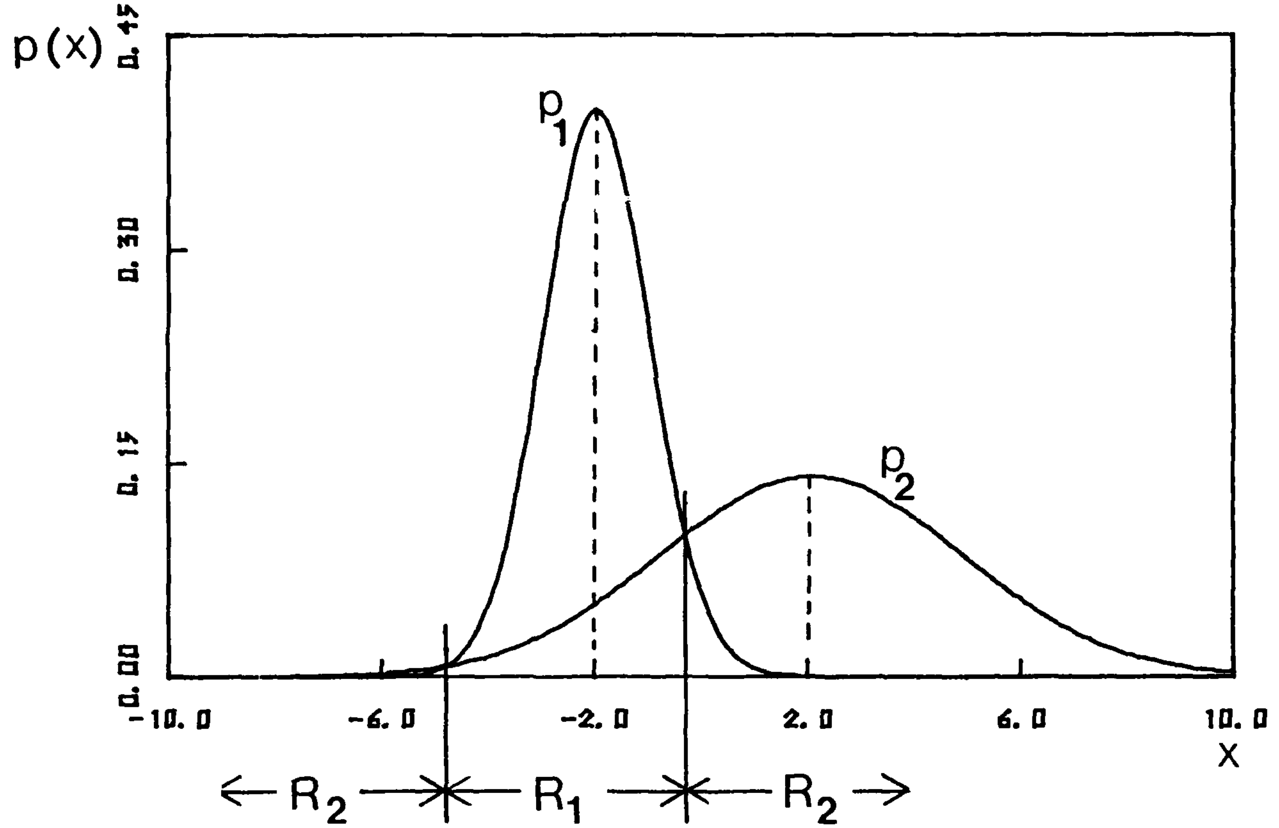
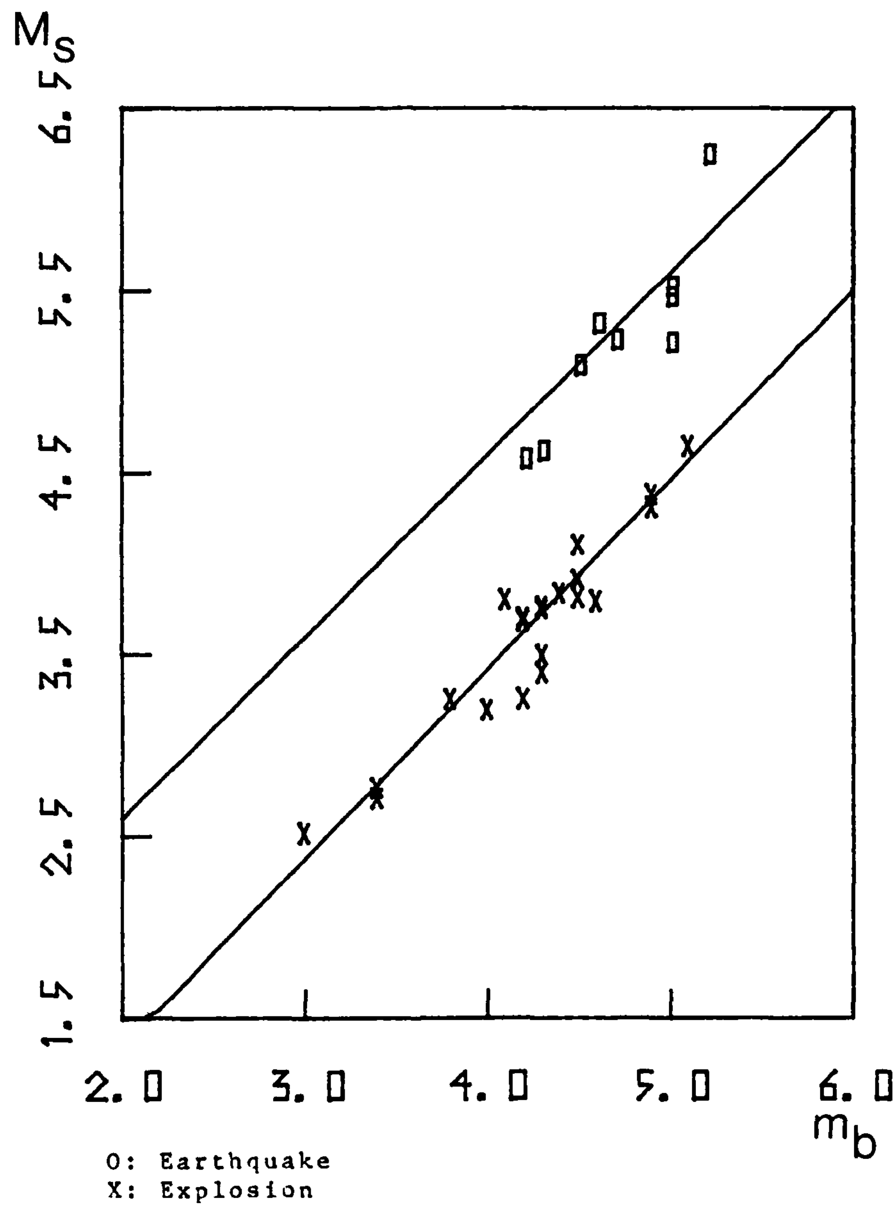
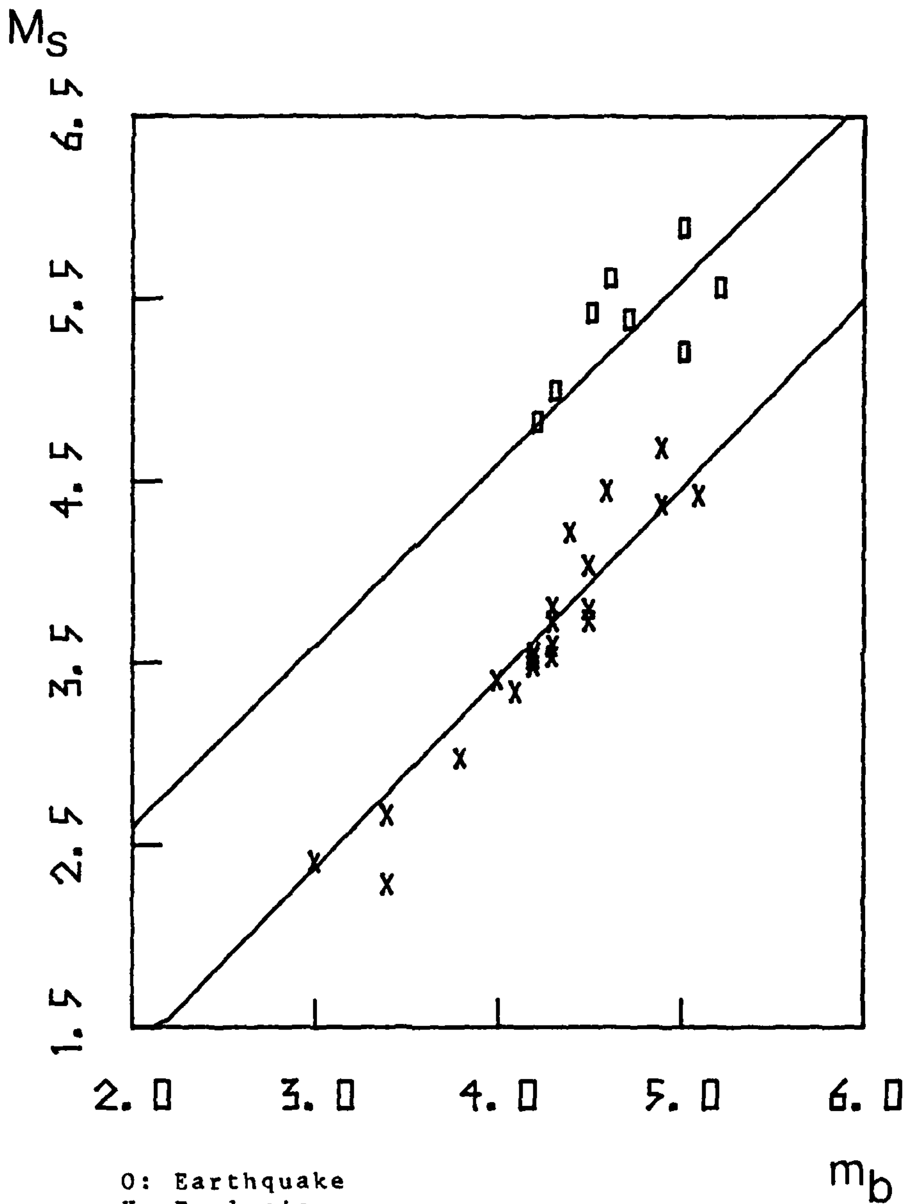


FIGURE (III - 2) Schematic representation of the classification regions in the two-group univariate quadratic discriminant analysis problem. The likelihood functions for the two groups are denoted by p_1 and p_2, the classification regions by R_1 and R_2.



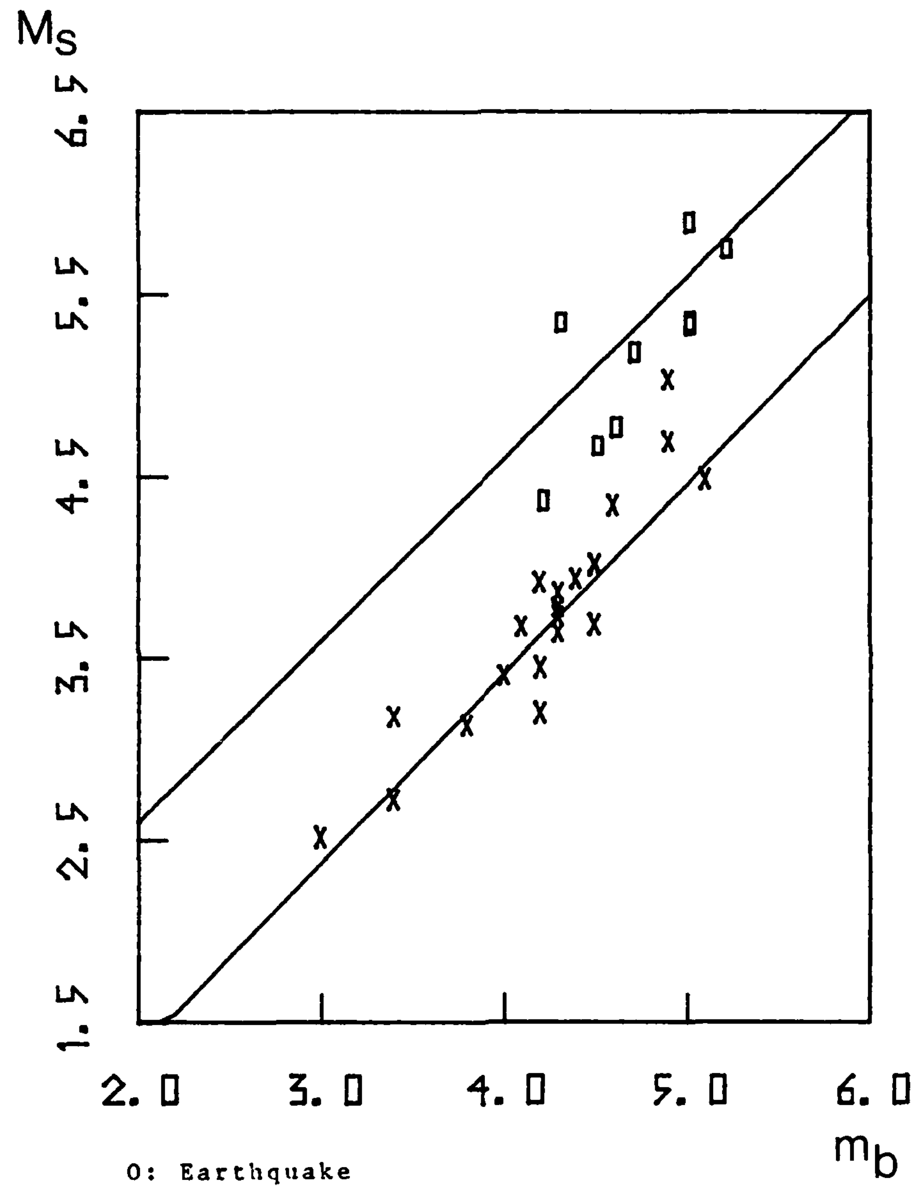
O: Earthquake
X: Explosion

FIGURE (IV - 1) M_s vs. m_b for M_s set 1.



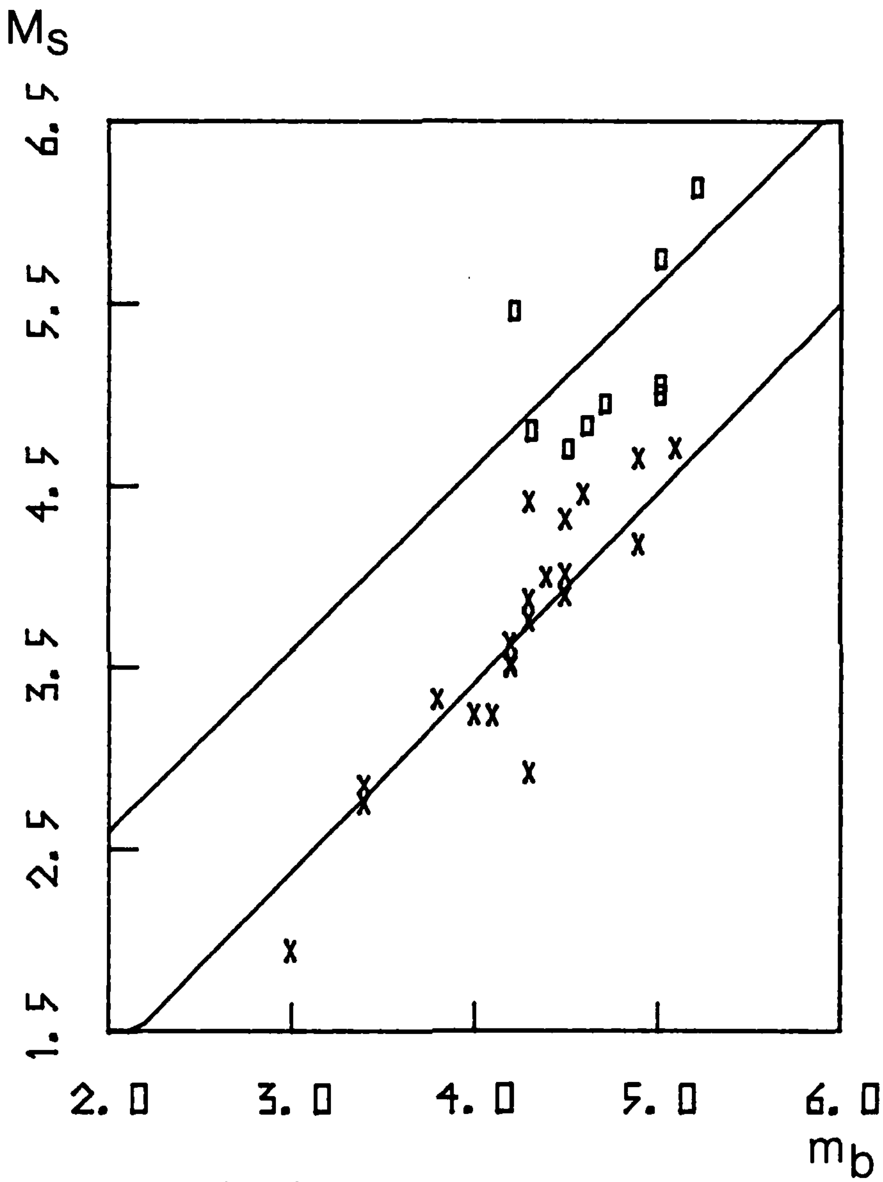
O: Earthquake
X: Explosion

FIGURE (IV - 2) M_s vs. m_b for M_s set 2.



O: Earthquake
X: Explosion

FIGURE (IV - 3) M_s vs. m_b for M_s set3.



O: Earthquake
X: Explosion

FIGURE (IV - 4) M_s vs. m_b for M_s set 4.

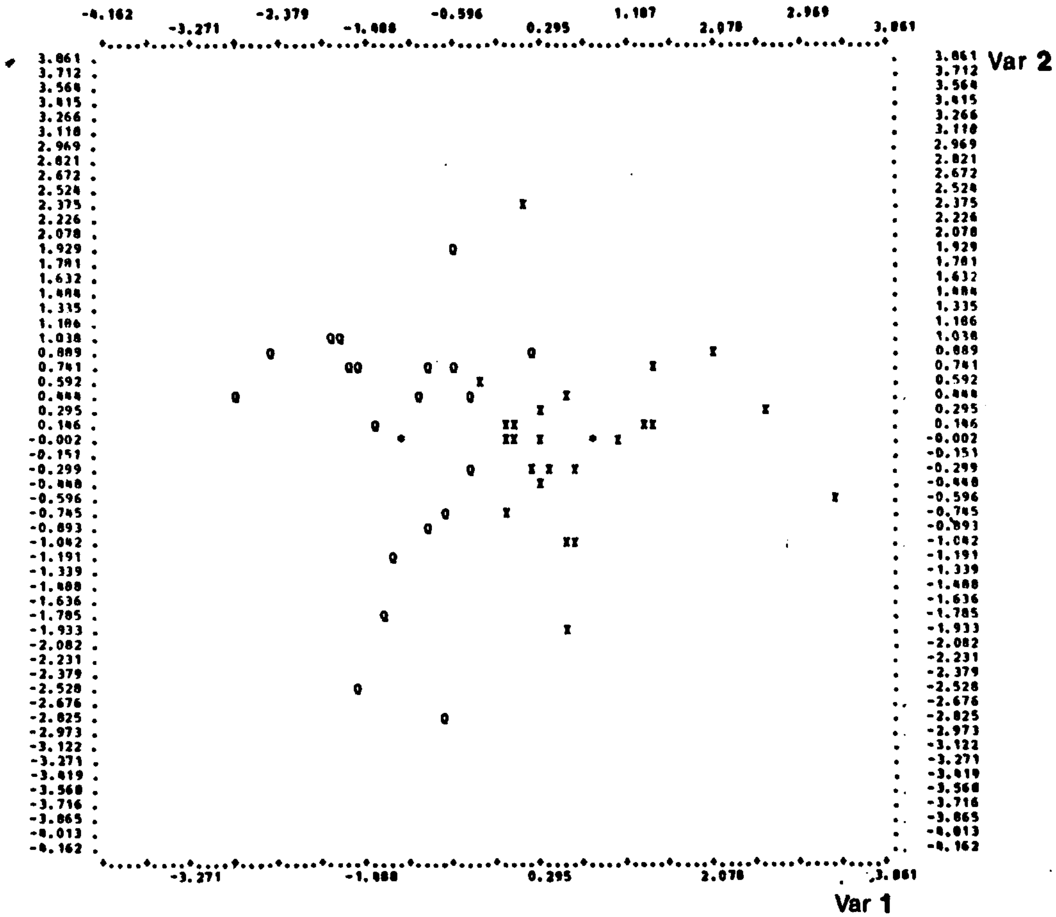


FIGURE IV-5: Scatterplot of 1st canonical variable (Var. 1) vs. 2nd canonical variable (Var. 2) for full set of 20 earthquakes and 27 explosions. 7 energy ratios were used. X denotes an explosion, Q denotes an earthquake.

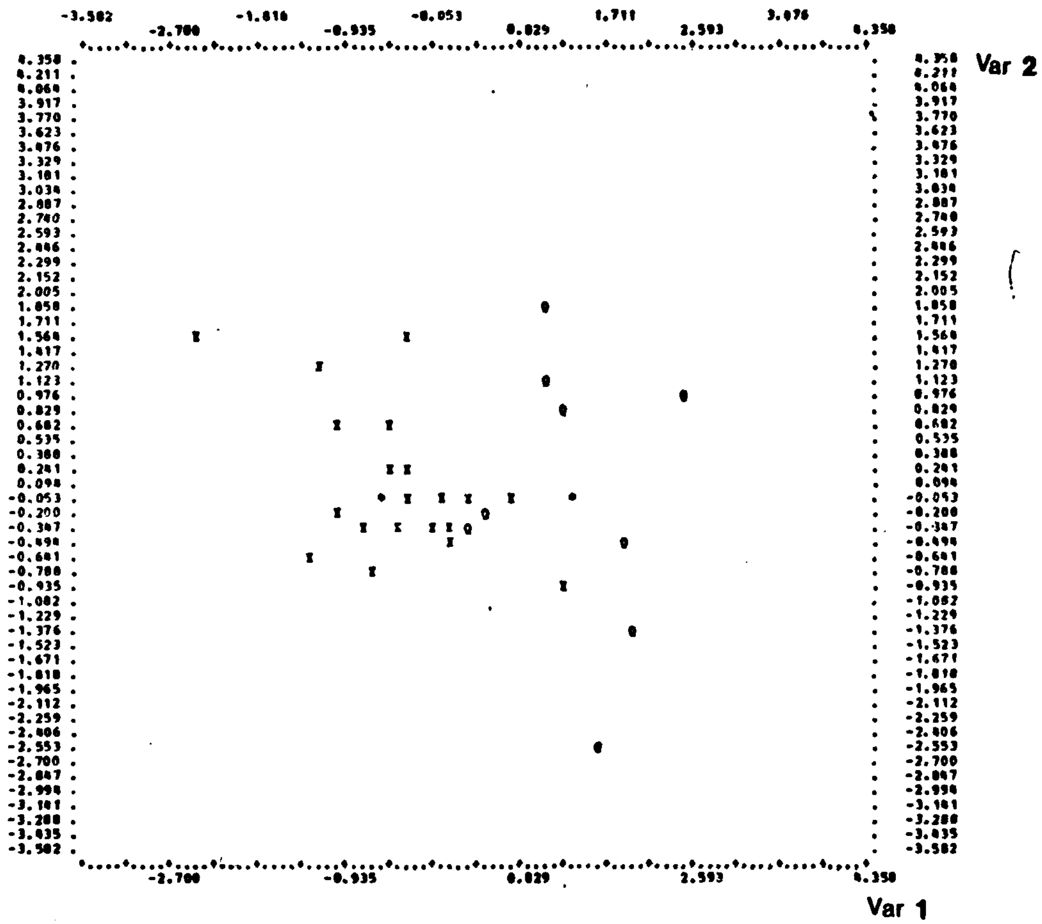


FIGURE IV-6: Scatterplot of 1st canonical variable (Var. 1) vs. 2nd canonical variable (Var. 2) for reduced set of 9 earthquakes and 21 explosions. 7 energy ratios were used. X denotes an explosion, Q denotes an earthquake.

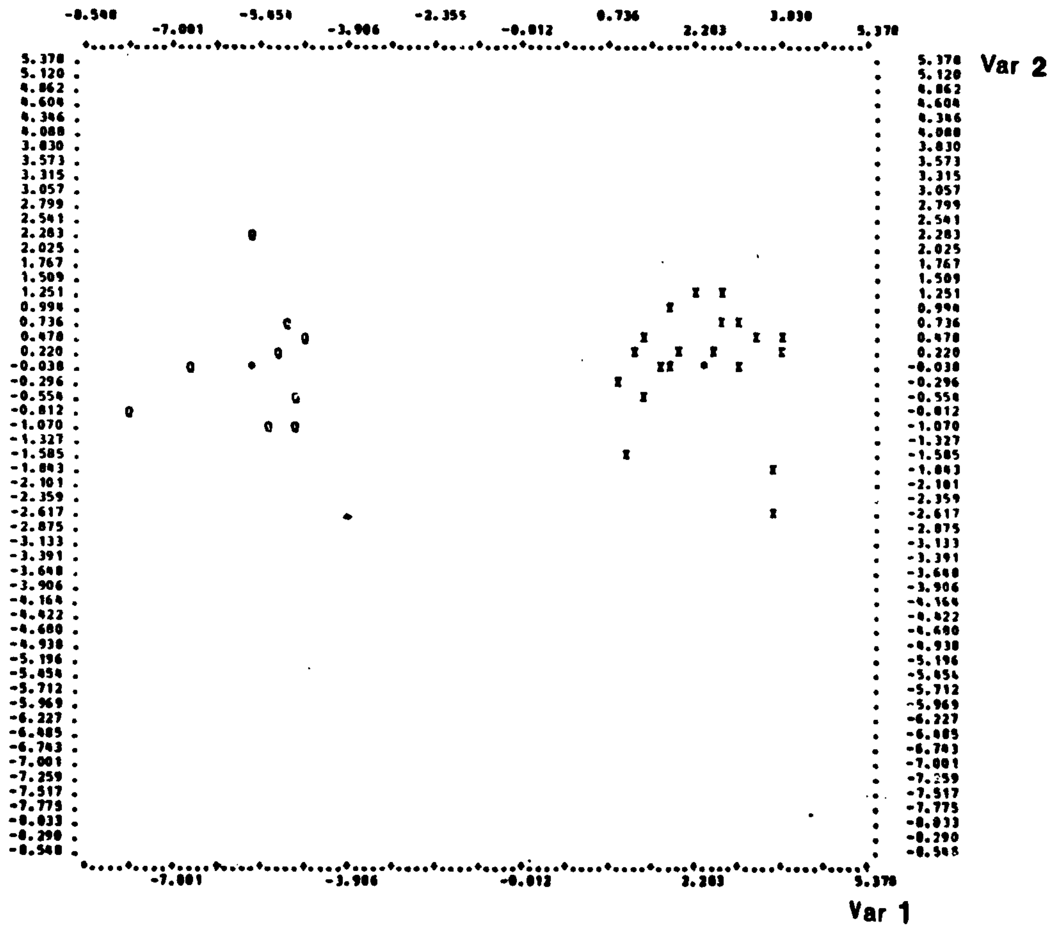


FIGURE IV-7: Scatterplot of 1st canonical variable (Var. 1) vs. 2nd canonical variable (Var. 2) for reduced set of 9 earthquakes and 21 explosions. 7 energy ratios and Ms/mb set 1 were used. X denotes an explosion, Q denotes an earthquake.

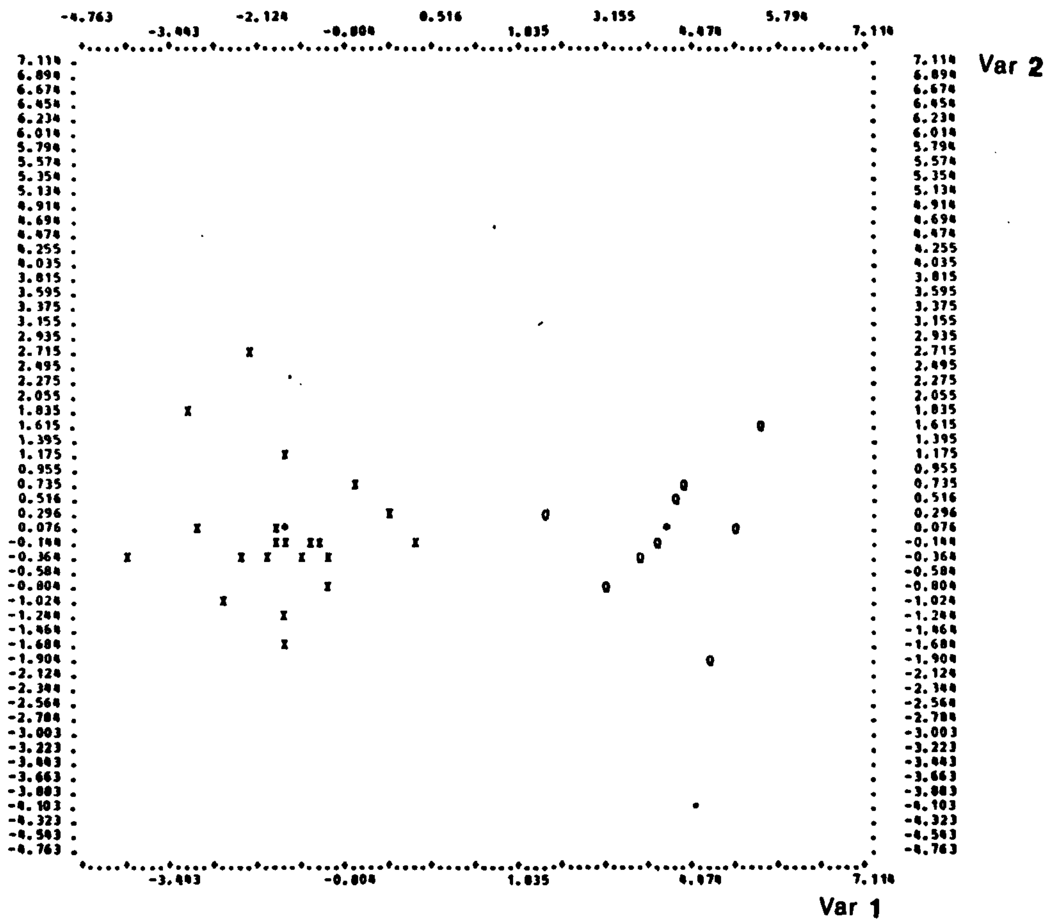


FIGURE IV-8: Scatterplot of 1st canonical variable (Var. 1) vs. 2nd canonical variable (Var. 2) for reduced set of 9 earthquakes and 21 explosions. 7 energy ratios and Ms/mb set 2 were used. X denotes an explosion, Q denotes an earthquake.

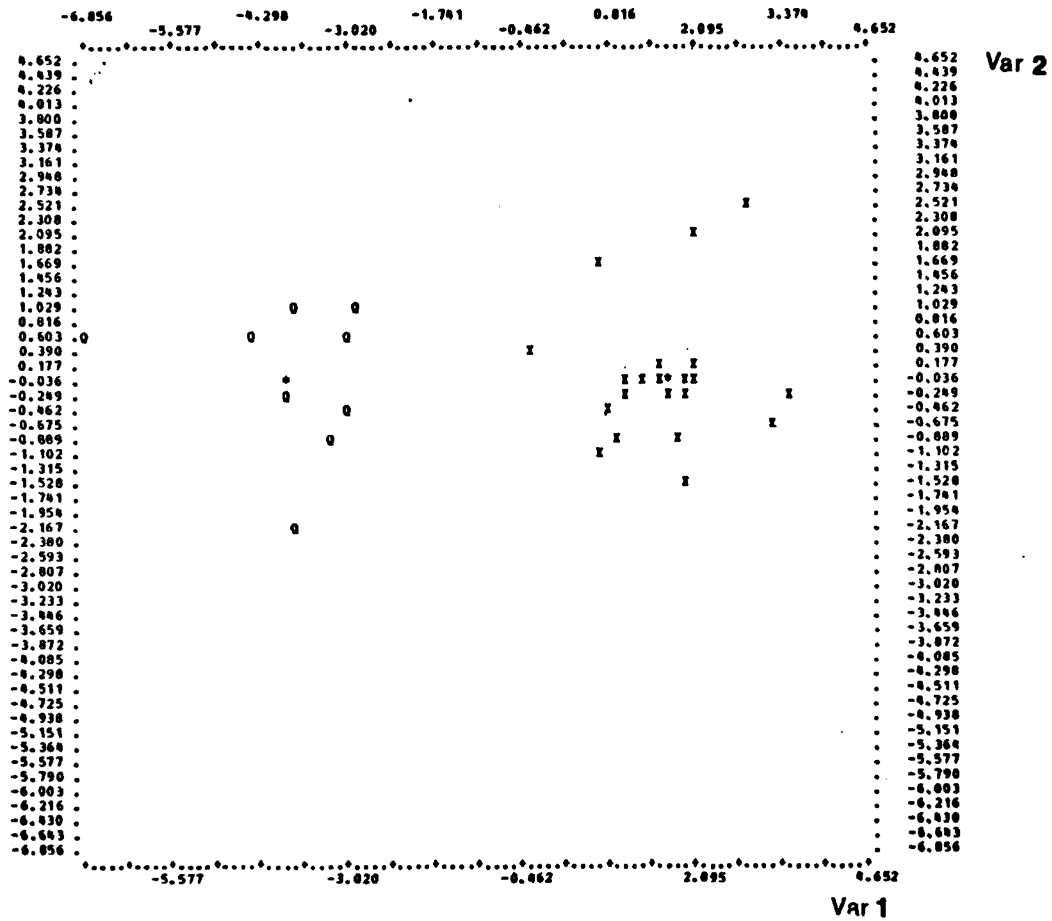


FIGURE IV-9: Scatterplot of 1st canonical variable (Var. 1) vs. 2nd canonical variable (Var. 2) for reduced set of 9 earthquakes and 21 explosions. 7 energy ratios and Ms/mb set 3 were used. X denotes an explosion, Q denotes an earthquake.

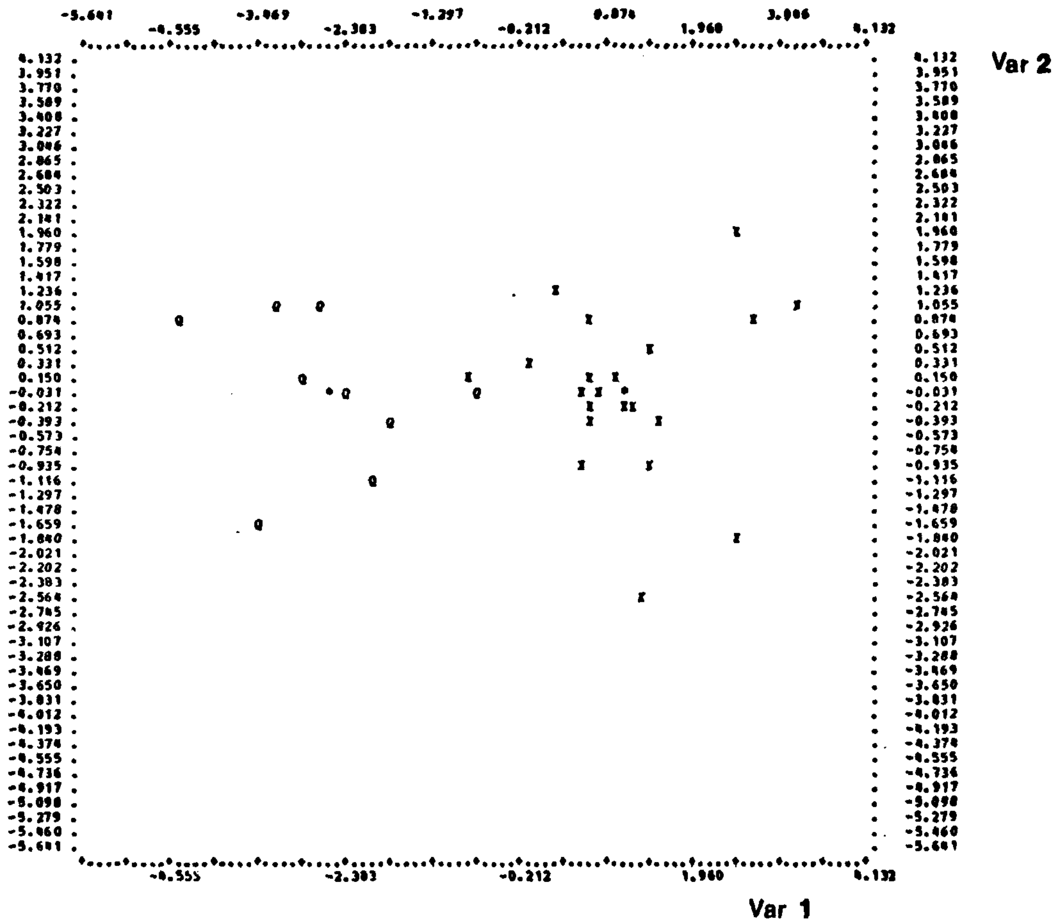


FIGURE IV-10: Scatterplot of 1st canonical variable (Var. 1) vs. 2nd canonical variable (Var. 2) for reduced set of 9 earthquakes and 21 explosions. 7 energy ratios and Ms/mb set 4 were used. X denotes an explosion, Q denotes an earthquake.

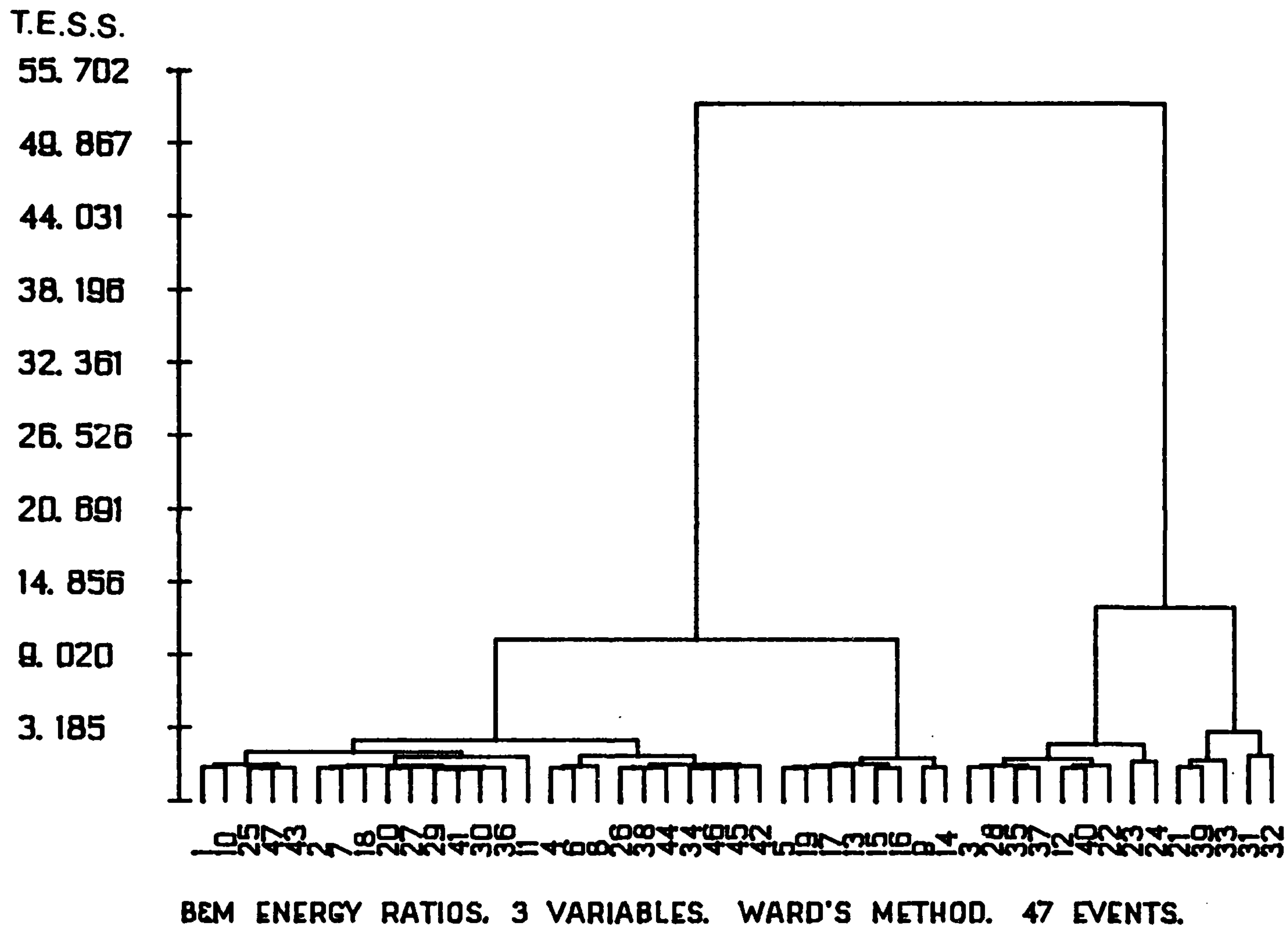


FIGURE (IV - 11) Cluster fusion history as a function of the total error sum of squares. Events 1-20 are earthquakes, and events 21-47 are explosions. See text for explanation of clustering method and variables used.

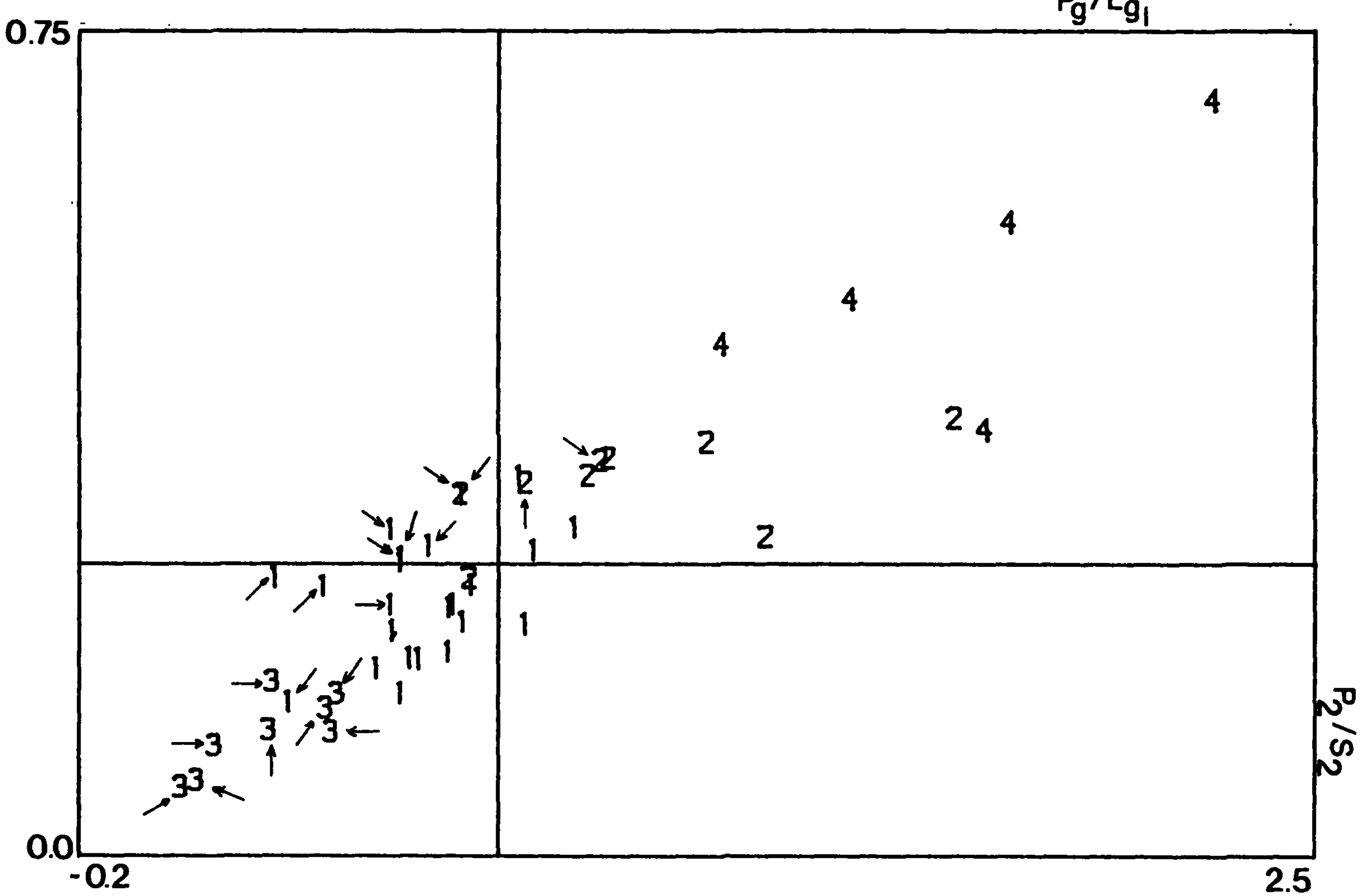


FIGURE (IV - 12) Scatterplot at four-cluster level of full set (20 eq., 27 ex.) of events. Ward's method was used. Arrows indicate earthquakes. Pg/Lg_1 is plotted horizontally, and P_2/S_2 is plotted vertically.

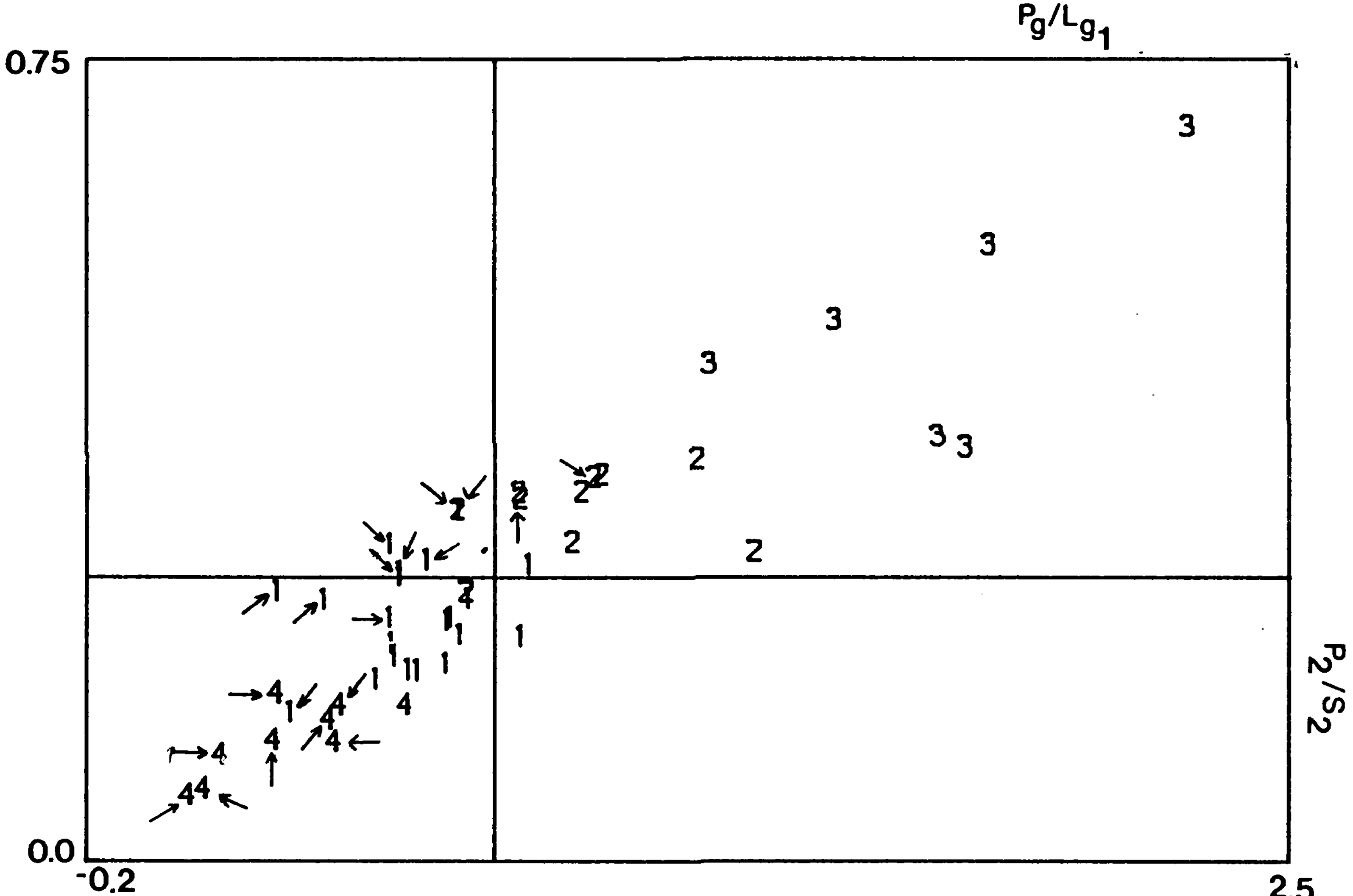
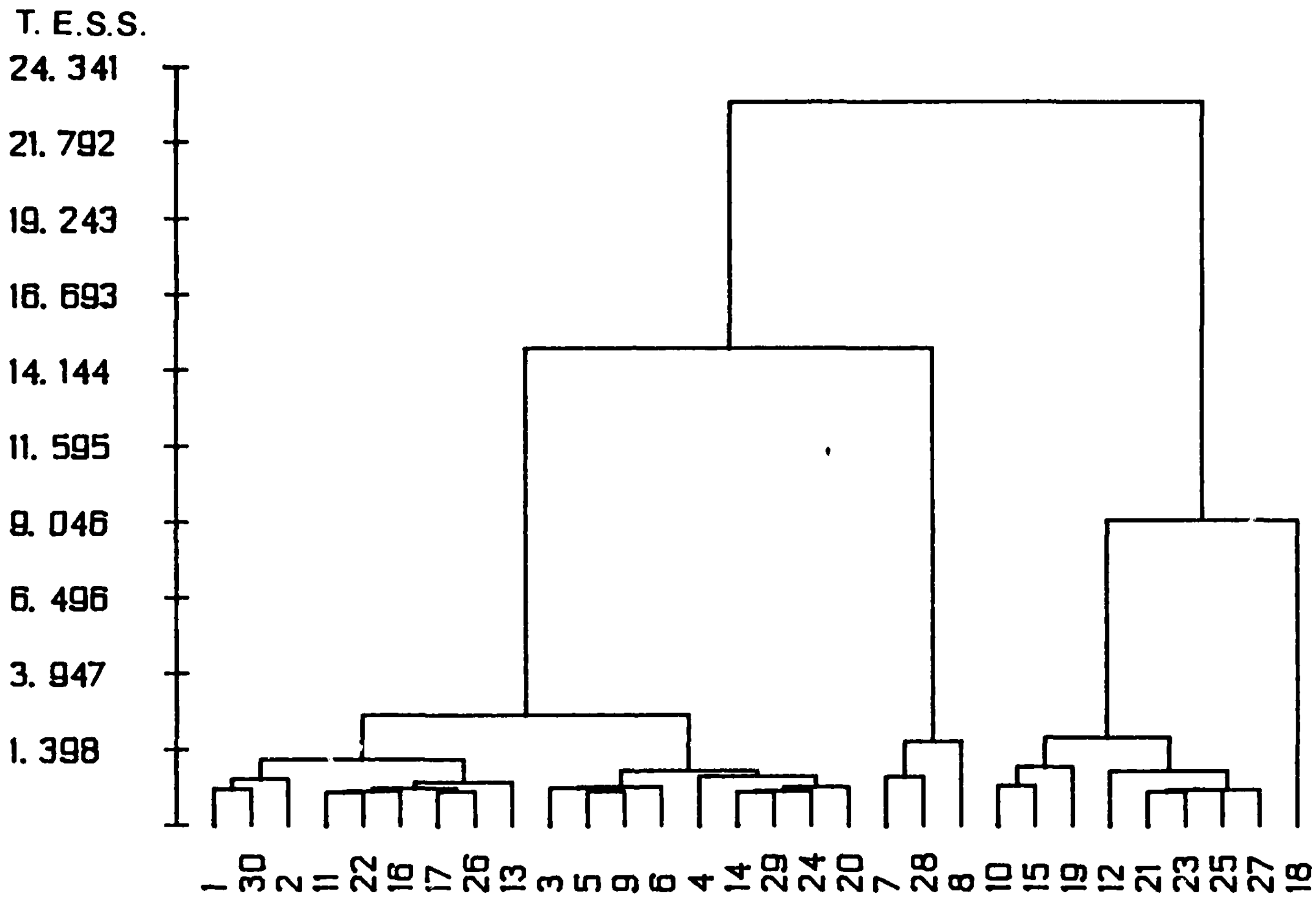


FIGURE (IV - 13) Scatterplot at four-cluster level of full set (20 eq., 27 ex.) of events. Procedure RELOCATE was used. Arrows indicate earthquakes. Pg/Lg_1 is plotted horizontally, and P_2/S_2 is plotted vertically.



B&M ENERGY RATIOS. 3 VARIABLES. WARD'S METHOD. 30 EVENTS.

FIGURE (IV - 14) Cluster fusion history as a function of the total error sum of squares. Events 1-9 are earthquakes, and events 21-47 are explosions. See text for explanation of clustering method and variables used.

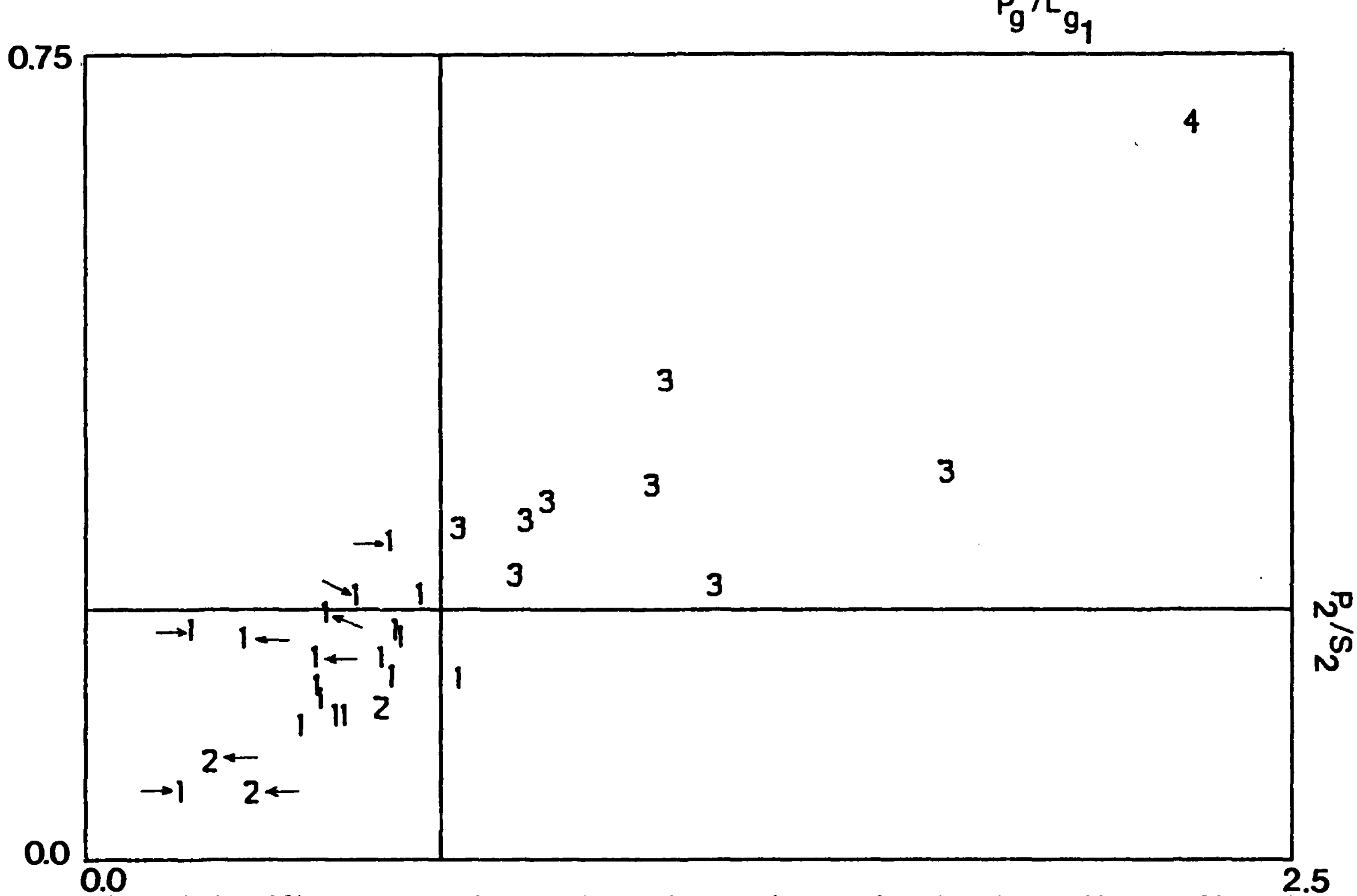


FIGURE (IV - 15) Scatterplot at four-cluster level of reduced set (9 eq., 21 ex.) of events. Ward's method was used. Arrows indicate earthquakes. Pg/Lg_1 is plotted horizontally, and P_2/S_2 is plotted vertically.

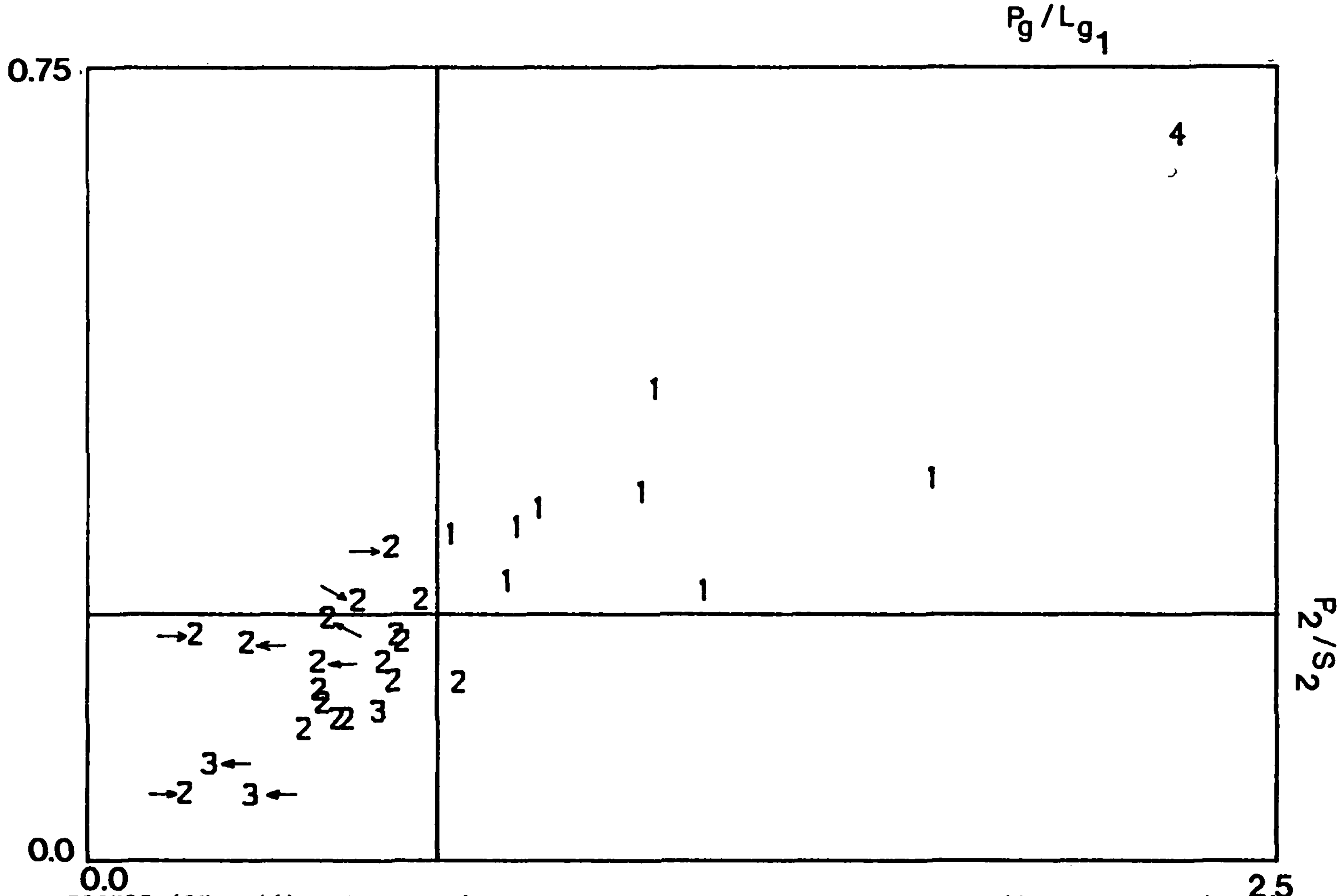
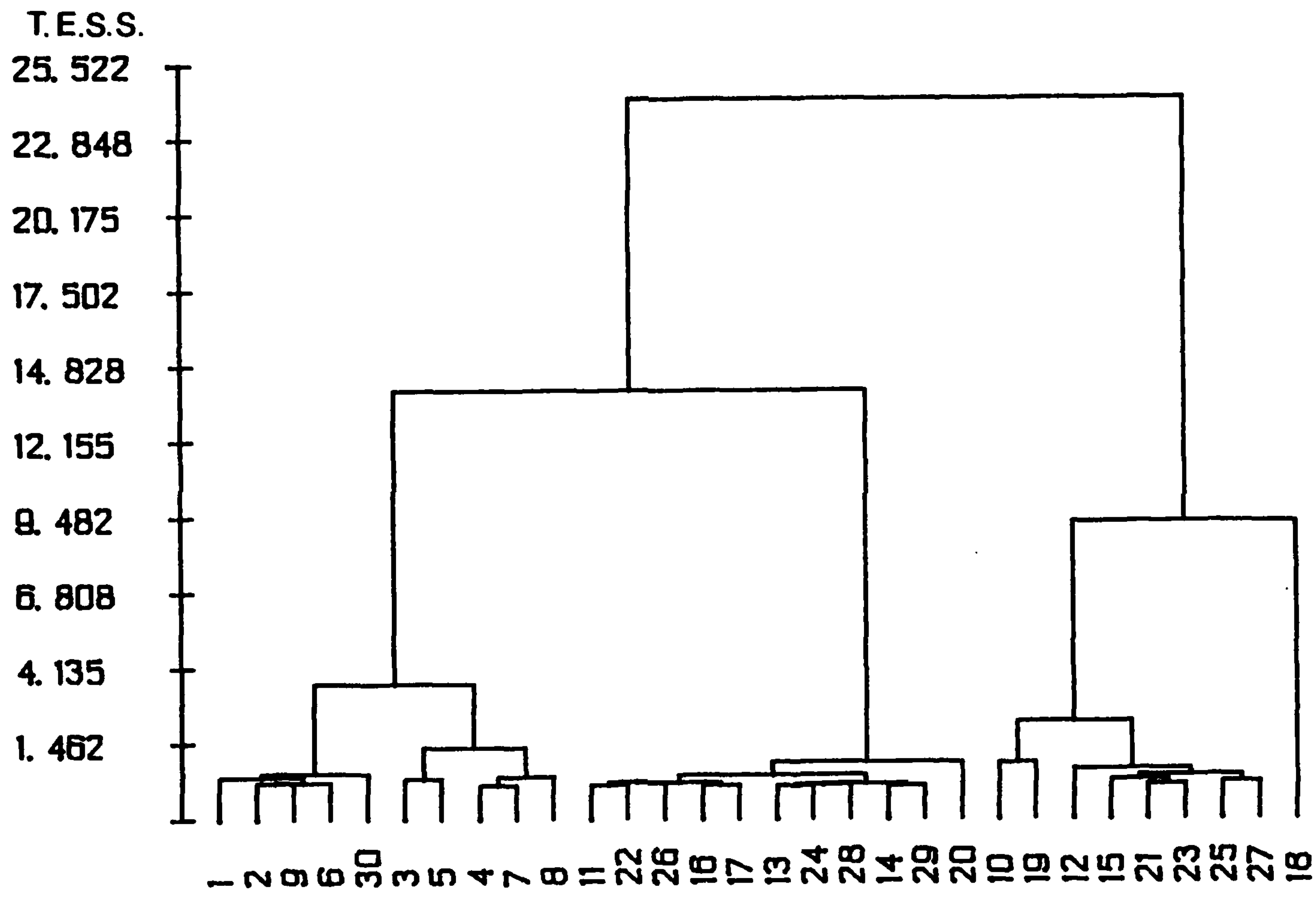


FIGURE (IV - 16) Scatterplot at four-cluster level of reduced (9 eq., 21 ex.) of events. Procedure RELOCATE was used. Arrows indicate earthquakes. Pg/Lg_1 is plotted horizontally, P_2/S_2 is plotted vertically.



B&M ENERGY RATIOS WITH MS / MB. SET 1. 3 VARIABLES. WARD'S METHOD.

FIGURE (IV - 17) Cluster fusion history as a function of the total error sum of squares. Events 1-9 are earthquakes, and events 10-30 are explosions. See text for explanation of clustering method and variables used.

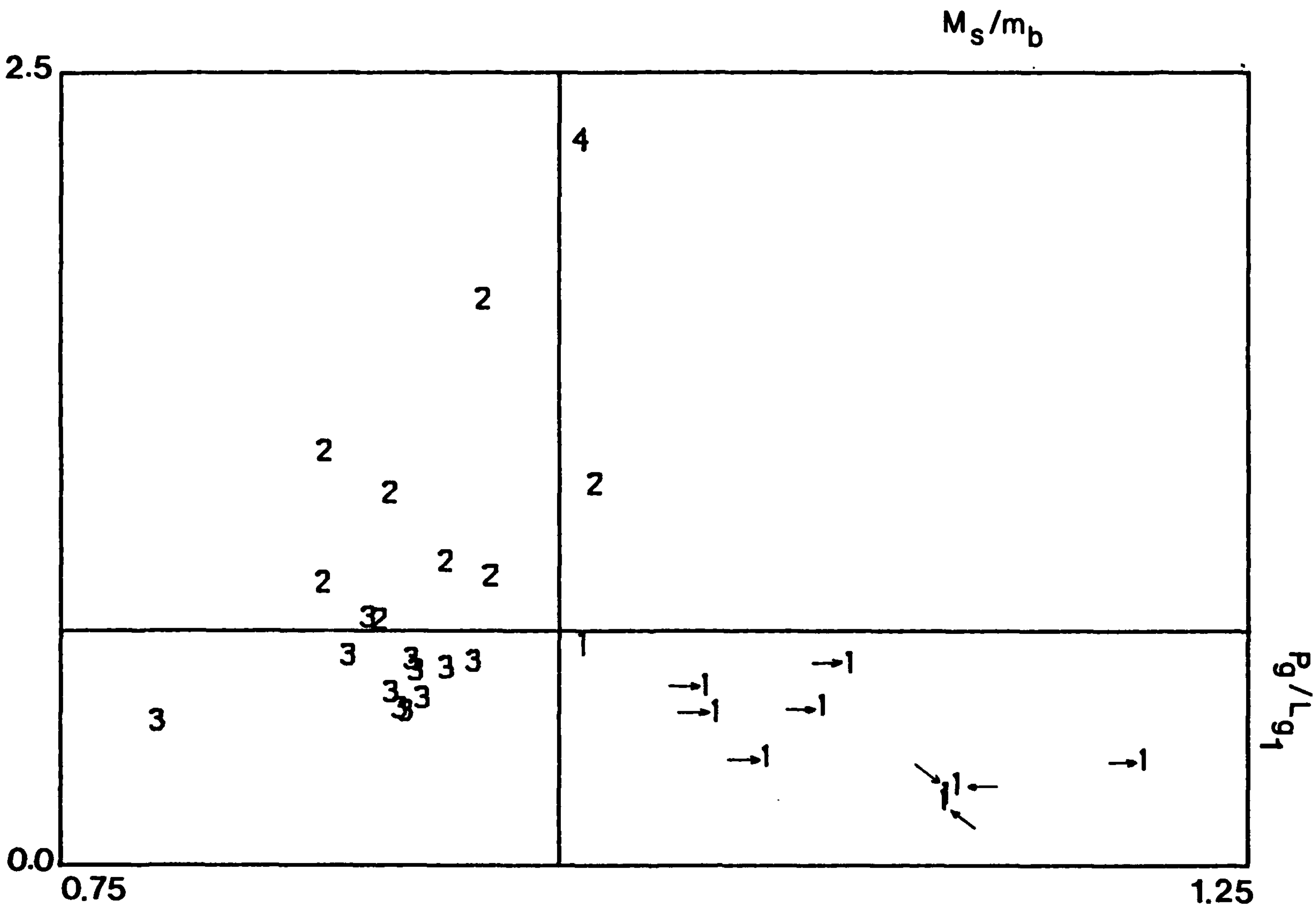


FIGURE (IV - 18) Scatterplot at four-cluster level of reduced set (9 eq., 21 ex.) of events with M_s/m_b included. Ward's method was used. Arrows indicate earthquakes. M_s/m_b is plotted horizontally, Pg/Lg_1 is plotted vertically.

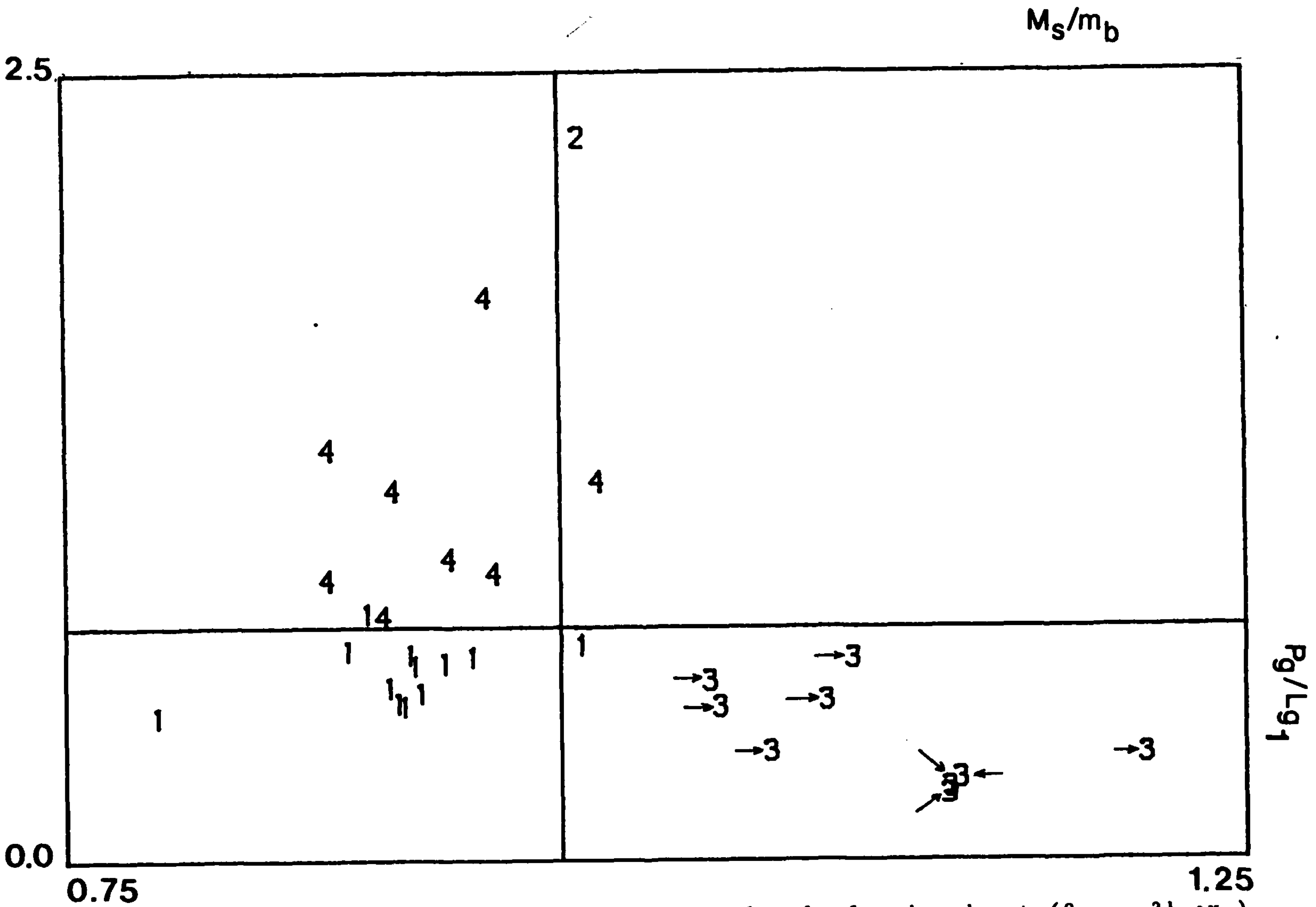


FIGURE (IV - 19) Scatterplot at four-cluster level of reduced set (9eq., 21 ex.) of events with M_s/m_b included. Procedure RELOCATE was used. Arrows indicate earthquakes. M_s/m_b is plotted horizontally, Pg/Lg_1 is plotted vertically.

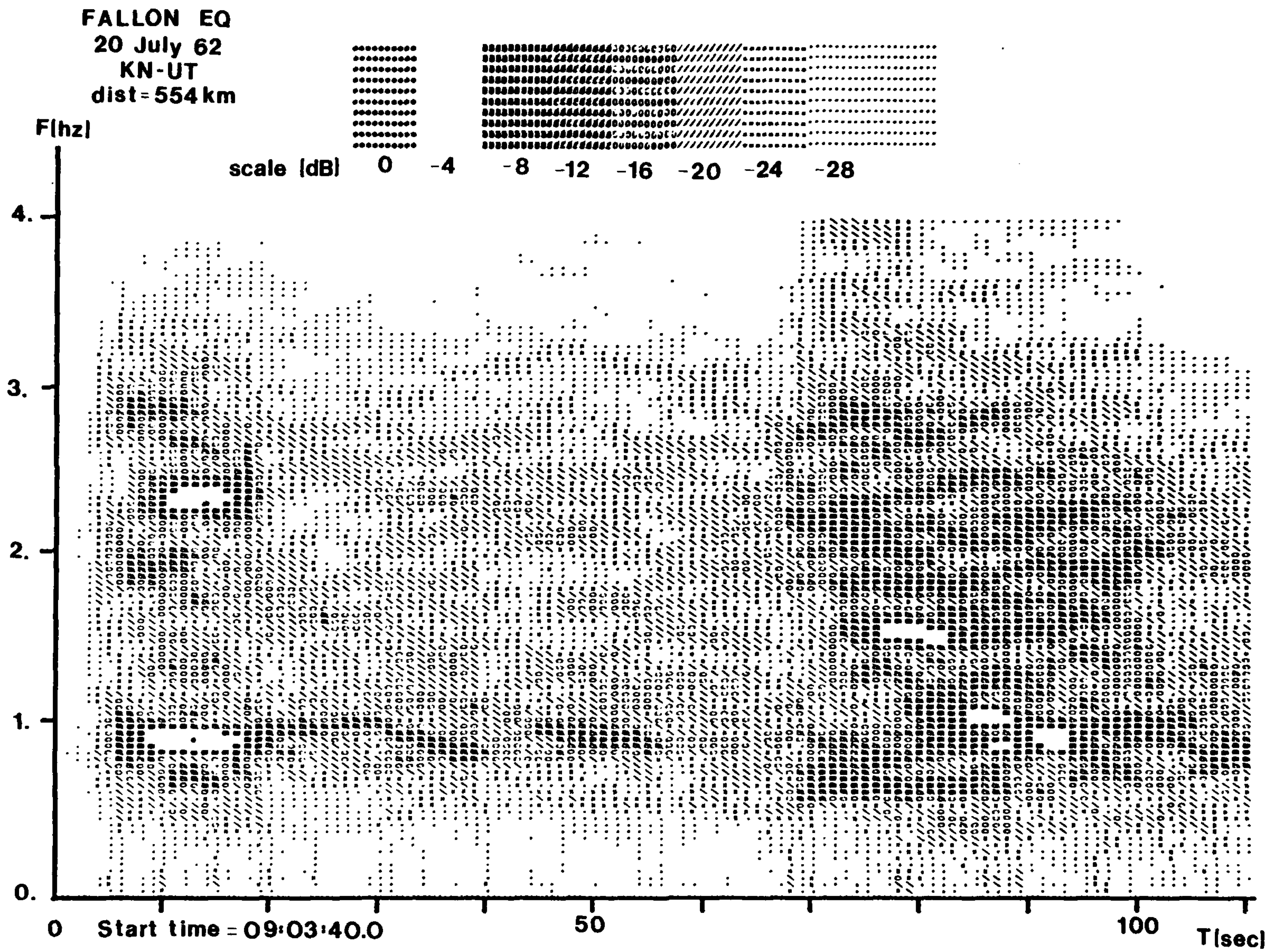


FIGURE (V-1): Seismoprint display of FALLON earthquake recorded at KN-UT.

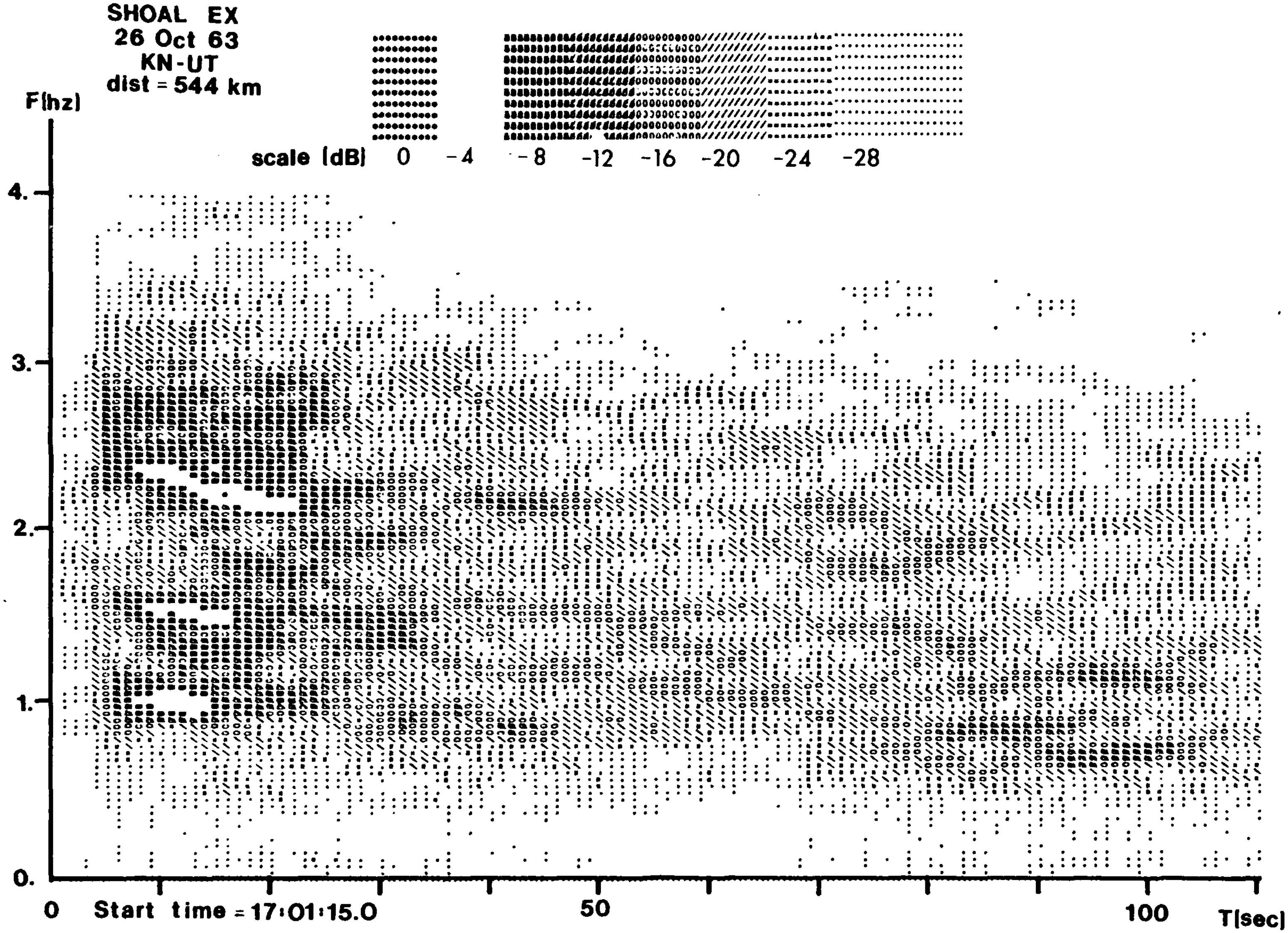


FIGURE (V-2): Seismoprint display of SHOAL explosion recorded at KN-UT.

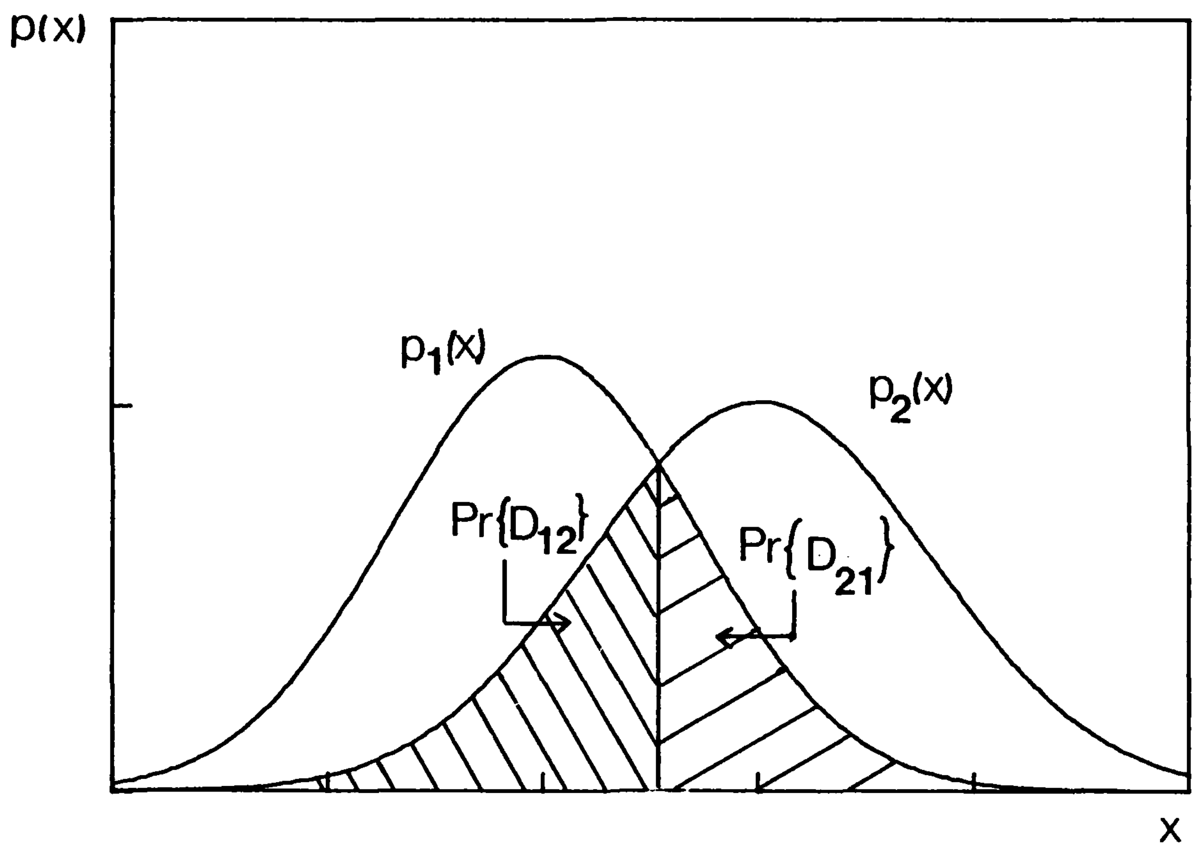


FIGURE (A - 1) Schematic representation of the error probabilities $Pr\{D_{12}\}$ and $Pr\{D_{21}\}$. The likelihood functions are denoted by $p_1(x)$ and $p_2(x)$.

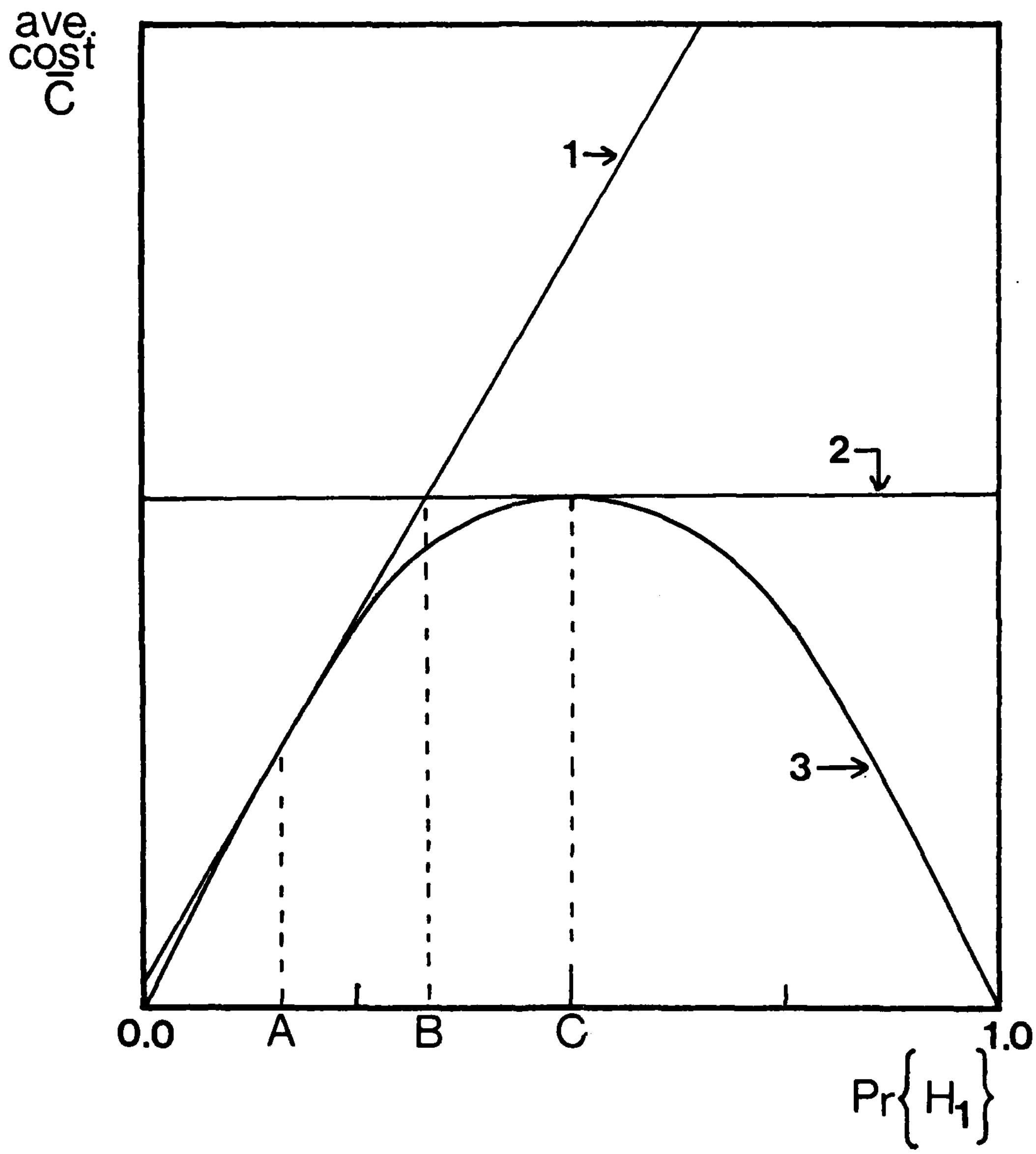


FIGURE (A - 2) Neymann-Pearson (1), minimax (2), and Bayes (3) costs as a function of $Pr\{H_1\}$. A and C are the values of $Pr\{H_1\}$ for which the Neymann-Pearson cost and minimax cost are tangent to the Bayes cost, and are the prior probabilities at which the two criteria are optimal. For $Pr\{H_1\} > B$, the Neymann-Pearson cost becomes highly suboptimal.

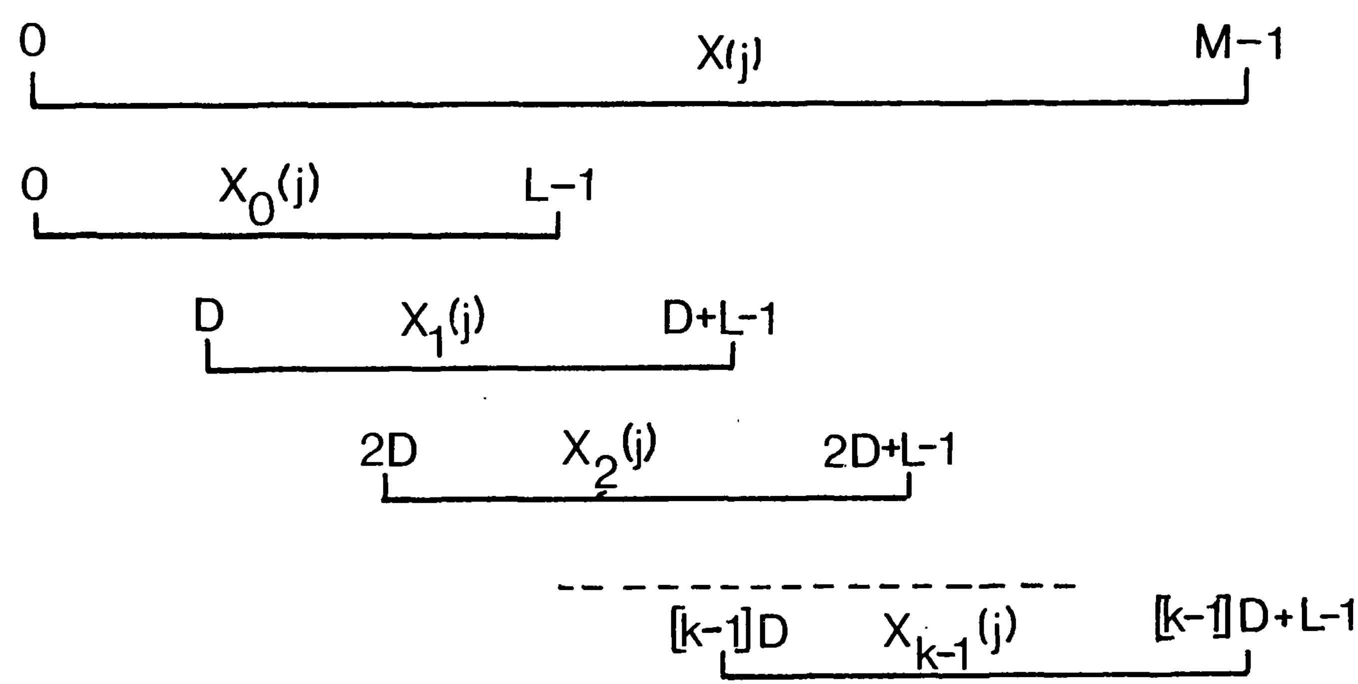


FIGURE (C - 1) Division of M sample record into K sections for obtaining spectral estimate by spectral averaging. Each segment is L samples long and is offset from the previous segment by D samples. Spectral estimates are obtained by calculating the modified periodogram for each segment, then averaging over the K periodograms.

TABLE IV-1
EVENT SUMMARY

Underground Nuclear Explosions

Event	Location lat.	Location long.	Date yr/mo/da	Origin Time hr m secs	m_b	Yld. (KT)	dist (Km)
AARDVARK	37.07 N	116.03 W	62/05/12	19 00 00.10	4.6	36.0	575
AGOUTI	37.05 N	116.03 W	62/01/18	18 00 00.13	4.5	5.8	530
ARMADILLO	37.04 N	116.04 W	62/02/09	16 30 00.13		6.5	536
CHINCHILLA II	37.05 N	116.04 W	62/03/31	18 00 00.10	3.8	< 20	527
CIMARRON	37.13 N	116.05 W	62/02/03	18 00 00.16		12.0	572
CODSAW	37.13 N	116.04 W	62/02/19	17 50 00.20	4.0	< 20	470
DANNY BOY	37.11 N	116.37 W	62/03/05	18 15 00.12	4.2	0.4	404
DESMOINES	37.22 N	116.16 W	62/06/13	21 00 00.12	4.5	< 20	511
DORMOUSE II	37.04 N	116.02 W	62/04/05	18 00 00.10	4.3	11.0	512
FISHER	37.05 N	116.03 W	61/12/03	23 04 59.63	4.3	13.0	481
GNOME	32.26 N	103.87 W	61/12/10	19 00 00.00		3.1	509
HARDHAT	37.23 N	116.06 W	62/02/15	18 00 00.10	4.9	5.9	559
HAYMAKER	37.04 N	116.04 W	62/06/27	18 00 00.12	4.9	46.0	621
MAD	37.13 N	116.05 W	61/12/13	18 00 00.16	3.4	0.5	475
MADISON	37.17 N	116.21 W	62/12/12	17 25 00.12	4.5	< 20	564
MARSHMALLOW	37.01 N	116.20 W	62/06/28	17 00 00.10	3.0	< 20	563
MISSISSIPPI	37.14 N	116.05 W	62/10/05	17 00 00.16	5.1	110.0	619
PACKRAT	37.05 N	116.04 W	62/06/06	17 00 00.10	4.2	< 20	435
PAMPAS	37.04 N	116.03 W	62/03/01	19 10 00.10		< 20	661
PASSAIC	37.12 N	116.04 W	62/04/06	18 00 00.20	4.3	< 20	618
PLATTE	37.22 N	116.16 W	62/04/14	18 00 00.10		1.9	410
SACRAMENTO	37.12 N	116.05 W	62/06/30	21 30 00.20	4.1	< 20	520
SMALL BOY	36.80 N	115.92 W	62/07/14	18 30 00.10	3.4		520
STILLWATER	37.13 N	116.05 W	62/02/08	18 00 00.10		3.1	456
STOAT	37.04 N	116.03 W	62/01/09	16 30 00.14	4.2		383
WICHITA	37.13 N	116.06 W	62/07/27	21 00 16.00	4.3	< 20	523
YORK	37.12 N	116.04 W	62/08/24	15 00 00.13	4.4	< 20	660

TABLE IV-1 (cont'd.)
EVENT SUMMARY

Earthquakes

Event	Location lat.	long.	Date yr/mo/da	Origin Time hr m secs	m_b	dist (Km)
Baja Ca	29.10 N	115.50 W	62/05/03	06 51 42.00	4.7	832
Baja Ca	29.10 N	115.50 W	62/05/03	13 48 24.00	5.0	1026
Boxelder Creek	39.90 N	104.60 W	62/12/05	13 48 00.40		746
Bridgeport	38.40 N	119.30 W	62/04/05	21 27 53.70	4.3	517
Cache Creek	41.80 N	111.80 W	62/08/30	13 25 28.70	4.5	1042
Cache Creek AS	40.70 N	112.00 W	62/09/15	16 04 29.00	4.6	685
Colona	38.20 N	107.60 W	62/02/05	14 45 51.10	4.2	656
Mont-Wyo Bdr.	45.00 N	110.20 W	62/07/15	11 59 21.90		794
Pierre S. Dak.	44.40 N	100.50 W	61/12/13	16 35 58.70		1116
Red Rock River	44.70 N	112.00 W	63/01/06	18 07 47.80		817
Sierra de Jua.	31.70 N	115.50 W	62/05/27	01 45 34.70	5.0	828
Teton County	43.60 N	110.80 W	62/10/06	09 28 17.40		635
W. Maryland	39.70 N	78.20 W	62/09/07	14 00 45.90		450
W. Vermont	30.80 N	73.10 W	62/04/10	14 30 46.40		1204
28 Jan 62	30.80 N	114.60 W	62/01/28	01 29 30.00		863
28 Jan 62	30.80 N	114.60 W	62/01/28	02 19 30.00		753
31 Jan 62	39.50 N	117.50 W	62/01/31	04 07 30.10		488
25 Apr 62	38.50 N	118.10 W	62/04/25	08 48 58.20		572
20 Aug 62	31.10 N	114.10 W	62/08/20	10 43 23.20	5.2	818
16 Nov 62	38.80 N	118.10 W	62/09/16	05 36 15.70	5.0	832

TABLE IV-2
BOOKER AND MITRONOVAS (1964) ENERGY RATIOS

Variable 1: P_1/S_1

Variable 2: P_2/S_2

Variable 3: Pg/B

Variable 4: Pg/Lg_1

Variable 5: $Pg/(Lg_1+Rg_1)$

Variable 6: Rg_1/Lg_1

Variable 7: Rg_2/Lg_2

Earthquakes

Clst.[1] Anal. Index	Dscrm.[1] Anal. Index	Event	Variable 1	2	3	4	5	6	7
1	1	Baja Ca	0.431	0.317	0.992	0.638	0.232	1.600	2.380
2	2	Baja Ca	0.341	0.265	1.204	0.568	0.142	2.560	3.380
3	3	Boxelder Creek	0.520	0.352	1.690	0.933	0.517	0.745	1.810
4	4	Bridgeport	0.319	0.232	0.927	0.231	0.180	0.290	0.775
5	5	Cache Creek	0.221	0.077	0.693	0.210	0.093	0.877	0.726
6	6	Cache Creek AS	0.324	0.224	0.819	0.338	0.210	0.613	1.330
7	7	Colona	0.333	0.248	1.074	0.506	0.485	0.336	0.526
8	8	Mont-Wyo Bdr.	0.302	0.281	0.815	0.486	0.368	0.740	1.315

TABLE IV-2 (cont'd.)

Clst.[1] Anal. Index	Dscrm.[1] Anal. Index	Event	Variable 1	2	3	4	5	6	7
9	9	Pierre S. Dak.	0.096	0.061	0.405	0.092	0.057	0.529	0.907
10	10	Red Rock River	0.490	0.336	1.190	0.765	0.457	0.710	1.070
11	11	Sierra de Jua.	0.156	0.105	1.690	0.261	0.092	1.640	6.750
12	12	Teton County	0.481	0.318	1.641	0.629	0.382	0.720	1.230
13	13	W. Maryland	0.046	0.026	0.841	0.053	0.030	1.520	1.210
14	14	W. Vermont	0.024	0.018	0.129	0.018	0.015	0.326	0.525
15	15	28 Jan 62	0.145	0.099	0.552	0.334	0.122	1.590	2.020
16	16	28 Jan 62	0.174	0.114	0.426	0.359	0.112	2.240	2.580
17	17	31 Jan 62	0.182	0.126	0.870	0.217	0.162	0.664	0.764
18	18	25 Apr 62	0.406	0.253	1.393	0.508	0.375	0.415	0.695
19	19	20 Aug 62	0.139	0.075	0.844	0.346	0.114	1.833	3.620
20	20	16 Nov 62	0.291	0.205	1.341	0.486	0.302	0.621	1.190

Underground Nuclear Explosions

Clst.[1] Anal. Index	Dscrm.[1] Anal. Index	Event	Variable 1	2	3	4	5	6	7
21	1	AARDVARK	0.636	0.471	2.550	1.194	0.620	0.907	1.370
22	2	AGOUTI	0.388	0.232	1.900	0.648	0.357	1.050	1.520
23	3	ARMADILLO	0.667	0.395	1.770	1.704	0.734	1.640	2.120
24	4	CHINCHILLA II	0.475	0.274	1.440	1.293	0.529	1.250	2.190
25	5	CIMARRON	0.479	0.260	1.320	0.796	0.352	1.500	1.852
26	6	CODSAW	0.312	0.150	0.840	0.526	0.267	1.000	2.270
27	7	DANNY BOY	0.315	0.178	1.269	0.488	0.252	0.972	1.500
28	8	DESMOINES	0.660	0.371	1.670	1.165	0.580	1.040	1.570
29	9	DORMOUSE II	0.442	0.205	1.380	0.620	0.293	1.670	1.450

TABLE IV-2 (cont'd.)

Clst.[1] Anal. Index	Dscrm.[1] Anal. Index	Event	Variable 1	2	3	4	5	6	7
30	10	FISHER	0.332	0.187	1.370	0.642	0.278	1.530	1.880
31	11	GNOME	0.897	0.596	1.960	1.819	1.114	0.611	2.860
32	12	HARDHAT	1.150	0.720	2.680	2.263	1.340	0.700	1.090
33	13	HAYMAKER	0.969	0.384	2.631	1.766	1.000	0.874	1.065
34	14	MAD	0.250	0.140	1.160	0.455	0.219	0.893	0.657
35	15	MADISON	0.485	0.337	2.010	0.906	0.500	0.915	1.720
36	16	MARSHMALLOW	0.396	0.225	1.380	0.660	0.330	1.070	1.330
37	17	MISSISSIPPI	0.527	0.355	1.940	0.951	0.488	0.965	1.386
38	18	PACKRAT	0.363	0.149	0.941	0.545	0.270	0.960	1.380
39	19	PAMPAS	0.804	0.517	2.650	1.473	0.900	0.629	1.030
40	20	PASSAIC	0.483	0.330	1.710	0.770	0.413	0.933	1.480
41	21	PLATTE	0.362	0.202	1.330	0.613	0.313	0.970	1.760
42	22	SACRAMENTO	0.353	0.185	1.070	0.777	0.314	1.110	1.717
43	23	SMALL BOY	0.399	0.284	1.390	0.885	0.434	1.170	1.330
44	24	STILLWATER	0.261	0.114	0.790	0.506	0.192	1.675	2.820
45	25	STOAT	0.400	0.157	1.130	0.612	0.257	2.070	5.590
46	26	WICHITA	0.297	0.165	1.150	0.495	0.248	0.935	1.360
47	27	YORK	0.422	0.266	1.310	0.699	0.346	1.040	1.840

1) The cluster analysis and discriminant analysis indices given in columns 1 and 2 are the numbers by which these events are referred to in the two analyses.

TABLE IV-3
ESTIMATED M_s AND M_s/m_b RATIOS

M_s and M_s/m_b set 1: $Var(M_s)$ = 0.0297 (explosions), 0.0594 (earthquakes)

" " " set 2: $Var(M_s)$ = 0.0446 (explosions), 0.0891 (earthquakes)

" " " set 3: $Var(M_s)$ = 0.0594 (explosions), 0.1188 (earthquakes)

" " set 4: $Var(M_s)$ = 0.0891 (explosions), 0.1792 (earthquakes)

(NOTE: 0.0297 is the explosion M_s variance given by von Seggern (1972).)

Earthquakes

Clst.[1] Anal. Index	Dscrm.[1] Anal. Index	Event	m_b	Set no.	1	2	3	4
					M_s estimates and M_s/m_b ratios			
1	1	Baja Ca	4.70	M_s	5.23	5.39	5.18	4.95
				M_s/m_b	1.11	1.15	1.10	1.05
2	2	Baja Ca	5.00	M_s	5.46	5.22	5.33	5.00
				M_s/m_b	1.09	1.04	1.07	1.00
3	3	Bridgeport	4.30	M_s	4.62	5.00	5.35	4.80
				M_s/m_b	1.07	1.16	1.24	1.12
4	4	Cache Creek	4.50	M_s	5.09	5.42	4.67	4.70
				M_s/m_b	1.13	1.21	1.04	1.04

TABLE IV-3 (cont'd.)

Clst.[1] Anal. Index	Dscrm.[1] Anal. Index	Event	m_b	Set no.	1	2	3	4
					M_s estimates and M_s/m_b ratios			
5	5	Cache Creek AS	4.60	M_s	5.32	5.62	4.77	4.83
				M_s/m_b	1.16	1.22	1.04	1.05
6	6	Colona	4.20	M_s	4.58	4.83	4.36	5.46
				M_s/m_b	1.09	1.15	1.04	1.30
7	7	Sierra de Jua.	5.00	M_s	5.22	5.90	5.90	5.75
				M_s/m_b	1.04	1.18	1.18	1.15
8	8	20 Aug 62	5.20	M_s	6.25	5.57	5.75	6.14
				M_s/m_b	1.20	1.07	1.11	1.18
9	9	16 Nov 62	5.00	M_s	5.52	5.21	5.35	5.06
				M_s/m_b	1.10	1.04	1.07	1.01

Underground Nuclear Explosions

Clst.[1] Anal. Index	Dscrm.[1] Anal. Index	Event	m_b	Set no.	1	2	3	4
					M_s estimates and M_s/m_b ratios			
10	1	AARDVARK	4.60	M_s	3.79	4.45	4.34	4.46
				M_s/m_b	0.82	0.97	0.94	0.97
11	2	AGOUTI	4.50	M_s	3.91	3.80	4.02	4.32
				M_s/m_b	0.87	0.84	0.89	0.96

TABLE IV-3 (cont'd.)

Clst.[1] Anal. Index	Dscrm.[1] Anal. Index	Event	m_b	M_s estimates and M_s/m_b ratios — Set no.	1	2	3	4
12	3	CHINCHILLA II	3.80	M_s	3.26	2.98	3.14	3.33
				M_s/m_b	0.86	0.78	0.83	0.88
13	4	CODSAW	4.00	M_s	3.20	3.40	3.40	3.24
				M_s/m_b	0.80	0.85	0.85	0.81
14	5	DANNY BOY	4.20	M_s	3.70	3.53	3.20	3.63
				M_s/m_b	0.88	0.84	0.76	0.87
15	6	DESMOINES	4.50	M_s	4.10	3.73	3.69	3.89
				M_s/m_b	0.91	0.83	0.82	0.86
16	7	DORMOUSE II	4.30	M_s	3.76	3.73	3.86	3.75
				M_s/m_b	0.88	0.87	0.90	0.87
17	8	FISHER	4.30	M_s	3.40	3.81	3.78	3.87
				M_s/m_b	0.79	0.89	0.88	0.90
18	9	HARDHAT	4.90	M_s	4.31	4.69	4.69	4.18
				M_s/m_b	0.88	0.96	0.96	0.85
19	10	HAYMAKER	4.90	M_s	4.38	4.37	5.03	4.65
				M_s/m_b	0.89	0.89	1.03	0.95
20	11	MAD	3.40	M_s	2.71	2.29	2.73	2.76
				M_s/m_b	0.80	0.67	0.80	0.81
21	12	MADISON	4.50	M_s	3.81	4.03	4.01	4.01
				M_s/m_b	0.85	0.90	0.89	0.89

TABLE IV-3 (cont'd.)

Clst.[1] Anal. Index	Dscrm.[1] Anal. Index	Event	m_b	Set no.	M_s estimates and M_s/m_b ratios 1	2	3	4
22	13	MARSHMALLOW	3.00	M_s	2.51	2.40	2.52	1.94
				M_s/m_b	0.84	0.80	0.84	0.65
23	14	MISSISSIPPI	5.10	M_s	4.65	4.42	4.48	4.71
				M_s/m_b	0.91	0.87	0.88	0.92
24	15	PACKRAT	4.20	M_s	3.68	3.48	3.45	3.52
				M_s/m_b	0.88	0.83	0.82	0.84
25	16	PASSAIC	4.30	M_s	3.49	3.53	3.73	4.41
				M_s/m_b	0.81	0.82	0.87	1.03
26	17	SACRAMENTO	4.10	M_s	3.80	3.34	3.68	3.23
				M_s/m_b	0.93	0.81	0.90	0.79
27	18	SMALL BOY	3.40	M_s	2.77	2.66	3.18	2.85
				M_s/m_b	0.81	0.78	0.94	0.84
28	19	STOAT	4.20	M_s	3.26	3.56	3.91	3.50
				M_s/m_b	0.78	0.85	0.93	0.83
29	20	WICHITA	4.30	M_s	3.74	3.59	3.64	2.92
				M_s/m_b	0.87	0.84	0.85	0.68
30	21	YORK	4.40	M_s	3.84	4.22	3.93	4.00
				M_s/m_b	0.87	0.96	0.89	0.91

1) The cluster analysis and discriminant analysis indices given in columns 1 and 2 are the numbers by which these events are referred to in the two analyses.

TABLE IV-4
VELOCITY WINDOWS USED IN FORMING ENERGY RATIOS

P_1	:	First arrival to 4.6 km/sec
P_2	:	First arrival to 4.9 km/sec
Pg	:	6.2 km/sec to 4.9 km/sec
S_1	:	4.6 km/sec to 2.5 km/sec
S_2	:	4.9 km/sec to 2.0 km/sec
B	:	4.9 km/sec to 3.6 km/sec
Lg_1	:	3.6 km/sec to 3.2 km/sec
Lg_2	:	3.6 km/sec to 3.2 km/sec[1]
Rg_1	:	3.2 km/sec to 2.8 km/sec
Rg_2	:	3.2 km/sec to 2.8 km/sec[1]

Ratio 1:	P_1/S_1
Ratio 2:	P_2/S_2
Ratio 3:	Pg/B
Ratio 4:	Pg/Lg_1
Ratio 5:	$Pg/(Lg_1+Rg_1)$
Ratio 6:	Rg_1/Lg_1
Ratio 7:	Rg_2/Lg_2

1) The energy within the windows Lg_2 and Rg_2 was computed using only the short-period vertical and radial components. For the rest of the windows, all three short-period components were used.

TABLE IV-5

BMDO7M stepwise linear discriminant analysis summary.
Full event set (20 earthquakes, 21 explosions).
7 energy ratios

```
STEP NUMBER        0
VARIABLE ENTERED

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1   45

  1  15.5225          2   6.9111          3  16.5384          4  20.5477

  5  12.3392          6   0.3182          7   0.0149


STEP NUMBER        1
VARIABLE ENTERED   4

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1   45

  4  20.5477

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1   44

  1   0.5601          2   7.1432          3   0.6954          5   4.8120

  6   1.0277          7   0.0614

               NUMBER OF CASES CLASSIFIED INTO GROUP -
               Q EQ    X EX
GROUP
  Q EQ          18       2
  X EX          12      15


STEP NUMBER        2
VARIABLE ENTERED   2

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1   44

  2   7.1432          4  20.5779

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1   43

  1   1.2729          3   4.0813          5   0.9350          6  -0.0000

  7   0.1433

               NUMBER OF CASES CLASSIFIED INTO GROUP -
               Q EQ    X EX
GROUP
  Q EQ          18       2
  X EX           6      21
```

TABLE IV-5 (cont'd.)

STEP NUMBER 3
VARIABLE ENTERED 3

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 43

2 10.8734 3 4.0813 4 14.3201

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 42

1 0.4096 5 1.3209 6 0.0000 7 0.4472

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	17	3
X EX	4	23

STEP NUMBER 4
VARIABLE ENTERED 5

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 42

2 5.3338 3 4.4233 4 11.4013 5 1.3209

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 41

1 1.0757 6 1.2486 7 1.5024

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	17	3
X EX	4	23

STEP NUMBER 5
VARIABLE ENTERED 7

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 41

2 5.7016 3 5.2738 4 12.9951 5 2.3773

7 1.5024

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 40

1 0.8992 6 0.2499

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	19	1
X EX	3	24

TABLE IV-5 (cont'd.)

STEP NUMBER 6
VARIABLE ENTERED 1

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 40

1	0.8992	2	6.5647	3	3.8146	4	8.4569
5	2.9858	7	1.3140				

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 39

6 0.3444

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	19	1
X EX	1	26

STEP NUMBER 7
VARIABLE ENTERED 6

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 39

1	0.9790	2	6.5264	3	3.5723	4	6.8842
5	2.8710	6	0.3444	7	0.3385		

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	19	1
X EX	1	26

TABLE IV-6

BMD07M stepwise linear discriminant analysis summary.
Reduced event set (9 earthquakes, 21 explosions).
7 energy ratios

STEP NUMBER 0
VARIABLE ENTERED

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 28

1	6.4082	2	2.5703	3	6.7749	4	9.2516
5	6.2126	6	0.0759	7	1.2271		

STEP NUMBER 1
VARIABLE ENTERED 4

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 28

4 9.2516

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 27

1	0.7125	2	3.4369	3	0.1491	5	1.2782
6	0.0040	7	0.3888				

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	8	1
X EX	7	14

STEP NUMBER 2
VARIABLE ENTERED 2

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 27

2 3.4369 4 10.0890

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 26

1	0.0043	3	0.9308	5	0.2950	6	0.1988
7	1.2887						

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	8	1
X EX	4	17

TABLE IV-6 (cont'd.)

```
STEP NUMBER         3
VARIABLE ENTERED    7

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1    26

  2    4.3256           4   10.8487           7    1.2887

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1    25

  1    0.0161           3    1.6678           5    1.1719           6    0.1983

             NUMBER OF CASES CLASSIFIED INTO GROUP -
             Q EQ    X EX
GROUP
  Q EQ         9       0
  X EX         2      19



STEP NUMBER         4
VARIABLE ENTERED    3

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1    25

  2    5.8284           3    1.6677           4    6.4002           7    2.0222

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1    24

  1    0.0545           5    2.1654           6    0.5121

             NUMBER OF CASES CLASSIFIED INTO GROUP -
             Q EQ    X EX
GROUP
  Q EQ         8       1
  X EX         3      18



STEP NUMBER         5
VARIABLE ENTERED    5

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1    24

  2    4.4566           3    2.6612           4    7.2318           5    2.1654

  7    3.7934

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1    23

  1    0.0652           6    0.0910

             NUMBER OF CASES CLASSIFIED INTO GROUP -
             Q EQ    X EX
GROUP
  Q EQ         9       0
  X EX         2      19
```

TABLE IV-6 (cont'd.)

```
STEP NUMBER        6
VARIABLE ENTERED   6

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1    23

  2    4.1686          3    2.5076          4    4.7277          5    1.6485

  6    0.0910          7    1.6784

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1    22

  1    0.1020

            NUMBER OF CASES CLASSIFIED INTO GROUP -
            Q EQ    X EX
GROUP
  Q EQ        8       1
  X EX        2      19


STEP NUMBER        7
VARIABLE ENTERED   1

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1    22

  1    0.1020          2    3.8732          3    2.3975          4    4.1866

  5    1.6523          6    0.1267          7    1.4518

            NUMBER OF CASES CLASSIFIED INTO GROUP -
            Q EQ    X EX
GROUP
  Q EQ        9       0
  X EX        2      19
```

TABLE IV-7

BMD07M stepwise linear discriminant analysis summary.
Reduced event set (9 earthquakes, 21 explosions).
7 energy ratios and M_s/m_b, set 1.

```
STEP NUMBER          0
VARIABLE ENTERED

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1   28

  1   6.4083         3   6.7749          5   6.2127          7   1.2271
  2   2.5704         4   9.2518          6   0.0760          8 214.5560



STEP NUMBER          1
VARIABLE ENTERED     8

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1   28

  8 214.5560

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1   27

  1   4.2278         2   1.4312          3   1.1862          4   5.3785

  5   3.7515         6   0.4996          7   3.8333

            NUMBER OF CASES CLASSIFIED INTO GROUP -
            Q EQ   X EX
GROUP
  Q EQ        9      0
  X EX        0     21



STEP NUMBER          2
VARIABLE ENTERED     4

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1   27

  4   5.3785         8 183.8242

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1   26

  1   0.0965         2   2.4298          3   0.9868          5   0.5922

  6   0.2065         7   2.7150

            NUMBER OF CASES CLASSIFIED INTO GROUP -
            Q EQ   X EX
GROUP
  Q EQ        9      0
  X EX        0     21
```

TABLE IV-7 (cont'd.)

STEP NUMBER 3
VARIABLE ENTERED 7

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 26

4	4.1540	7	2.7150	8	195.0327

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 25

1	0.6282	2	4.7524	3	0.5917	5	3.1773
6	0.9724						

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

STEP NUMBER 4
VARIABLE ENTERED 2

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 25

2	4.7524	4	8.8428	7	5.0509	8	191.8542

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 24

1	0.0238	3	0.0003	5	1.8229	6	0.5786

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

STEP NUMBER 5
VARIABLE ENTERED 5

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 24

2	3.2664	4	7.7013	5	1.8229	7	6.9314
8	189.9639						

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 23

1	0.4217	3	0.1677	6	0.0311

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

TABLE IV-7 (cont'd.)

```
STEP NUMBER         6
VARIABLE ENTERED    1

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1    23

  1    0.4217           2    3.5950          4    5.4019          5    2.1758

  7    6.6603           8  185.2673

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1    22

  3    0.1637           6    0.0990

             NUMBER OF CASES CLASSIFIED INTO GROUP -
             Q EQ    X EX
GROUP
  Q EQ         9       0
  X EX         0      21


STEP NUMBER         7
VARIABLE ENTERED    3

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1    22

  1    0.4067           2    3.5936          3    0.1637          4    5.3094

  5    2.2603           7    6.3169          8  158.6258

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1    21

  6    0.0879

             NUMBER OF CASES CLASSIFIED INTO GROUP -
             Q EQ    X EX
GROUP
  Q EQ         9       0
  X EX         0      21


STEP NUMBER         8
VARIABLE ENTERED    6

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1    21

  1    0.4524           3    0.1497          5    1.6573          7    3.0042
  2    3.4728           4    3.9643          6    0.0879          8  151.1451

             NUMBER OF CASES CLASSIFIED INTO GROUP -
             Q EQ    X EX
GROUP
  Q EQ         9       0
  X EX         0      21
```

TABLE IV-8

BMD07M stepwise linear discriminant analysis summary.
Reduced event set (9 earthquakes, 21 explosions).
7 energy ratios and M_s/m_b, set 2.

```
STEP NUMBER          0
VARIABLE ENTERED

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1   28

  1   6.4083        3   6.7749        5   6.2127        7    1.2271
  2   2.5704        4   9.2518        6   0.0760        8  115.5020



STEP NUMBER          1
VARIABLE ENTERED     8

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1   28

  8 115.5020

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1   27

  1   12.1118        2   8.5479        3   10.1405       4   8.2551

  5    9.3989        6   1.7685        7    0.3380

           NUMBER OF CASES CLASSIFIED INTO GROUP -
           Q EQ    X EX
GROUP
  Q EQ       9       0
  X EX       0      21



STEP NUMBER          2
VARIABLE ENTERED     1

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM     1   27

  1   12.1118        8 136.1186

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM     1   26

  2    0.0628        3   0.6956        4    1.0353       5   0.1996

  6    0.6064        7   0.0315

           NUMBER OF CASES CLASSIFIED INTO GROUP -
           Q EQ    X EX
GROUP
  Q EQ       9       0
  X EX       0      21
```

TABLE IV-8 (cont'd.)

```
STEP NUMBER          3
VARIABLE ENTERED     4

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1    26

  1    3.9929           4    1.0353           8 120.9839

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1    25

  2    0.0021           3    1.2139           5    0.0228           6    0.3319

  7    0.2162

             NUMBER OF CASES CLASSIFIED INTO GROUP -
             Q EQ    X EX
GROUP
  Q EQ         9       0
  X EX         0      21



STEP NUMBER          4
VARIABLE ENTERED     3

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1    25

  1    3.2136           3    1.2139           4    1.5474           8 121.5642

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1    24

  2    0.1019           5   -0.0002           6    0.3314           7    0.0544

             NUMBER OF CASES CLASSIFIED INTO GROUP -
             Q EQ    X EX
GROUP
  Q EQ         9       0
  X EX         0      21



STEP NUMBER          5
VARIABLE ENTERED     6

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1    24

  1    2.4207           3    1.1784           4    1.1714           6    0.3314

  8 118.4616

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1    23

  2    0.1690           5    0.4007           7    0.8938

             NUMBER OF CASES CLASSIFIED INTO GROUP -
             Q EQ    X EX
GROUP
  Q EQ         9       0
  X EX         0      21
```

TABLE IV-8 (cont'd.)

STEP NUMBER 6
VARIABLE ENTERED 7

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 23

1	3.0411	3	0.5367	4	1.3800	6	1.1689
7	0.8938	8	110.5479				

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 22

2	0.0344	5	0.5757

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

STEP NUMBER 7
VARIABLE ENTERED 5

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 22

1	3.5437	3	0.7102	4	0.0343	5	0.5757
6	1.7097	7	1.0514	8	100.5857		

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 21

2	0.0534

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

STEP NUMBER 8
VARIABLE ENTERED 2

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 21

1	3.3340	3	0.7310	5	0.5706	7	0.8444
2	0.0534	4	0.0182	6	1.6286	8	78.7544

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

TABLE IV-9

BMD07M stepwise linear discriminant analysis summary. Reduced event set (9 earthquakes, 21 explosions). 7 energy ratios and M_s/m_b, set 3.

STEP NUMBER 0
VARIABLE ENTERED

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 28

1	6.4083	3	6.7749	5	6.2127	7	1.2271
2	2.5704	4	9.2518	6	0.0760	8	74.7827

STEP NUMBER 1
VARIABLE ENTERED 8

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 28

8 74.7827

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 27

1	14.2533	2	7.2224	3	19.5475	4	14.0003
5	11.5038	6	0.0014	7	0.1908		

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	1	20

STEP NUMBER 2
VARIABLE ENTERED 3

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 27

3 19.5476 8 110.5786

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 26

1	0.6249	2	0.2939	4	0.4700	5	0.0675
6	0.2538	7	0.6005				

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

TABLE IV-9 (cont'd.)

STEP NUMBER 3
VARIABLE ENTERED 1

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 26

1	0.6249	3	4.0418	8	106.8292

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 25

2	3.2717	4	0.0001	5	0.7387	6	0.4122
7	1.2722						

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

STEP NUMBER 4
VARIABLE ENTERED 2

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 25

1	3.6276	2	3.2717	3	6.0983	8	104.8908

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 24

4	0.0173	5	0.4550	6	0.1527	7	0.5616

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

STEP NUMBER 5
VARIABLE ENTERED 7

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 24

1	3.7354	2	2.4308	3	4.3818	7	0.5616
8	97.4571						

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 23

4	0.0050	5	0.2534	6	0.0419

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

TABLE IV-9 (cont'd.)

STEP NUMBER 6
VARIABLE ENTERED 5

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 23

1	2.7436	2	2.2458	3	4.4862	5	0.2534
7	0.3548	8	94.6508				

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 22

4	0.0945	6	0.2211

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

STEP NUMBER 7
VARIABLE ENTERED 6

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 22

1	2.8315	2	2.0756	3	4.2000	5	0.4251
6	0.2211	7	0.5570	8	86.6240		

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 21

4	0.4590

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

STEP NUMBER 8
VARIABLE ENTERED 4

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 21

1	2.2404	3	4.0976	5	0.8714	7	0.5601
2	2.2193	4	0.4590	6	0.5819	8	68.0162

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

TABLE IV-10

BMD07M stepwise linear discriminant analysis summary.
Reduced event set (9 earthquakes, 21 explosions).
7 energy ratios and M_s/m_b, set 4.

STEP NUMBER 0
VARIABLE ENTERED

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 28

1	6.4083	3	6.7749	5	6.2127	7	1.2271
2	2.5704	4	9.2518	6	0.0760	8	43.6063

STEP NUMBER 1
VARIABLE ENTERED 8

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 28

8 43.6063

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 27

1	6.2522	2	4.1747	3	12.8523	4	7.8121
5	8.2655	6	0.7123	7	0.4637		

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	1	20

STEP NUMBER 2
VARIABLE ENTERED 3

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 27

3 12.8523 8 55.0613

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 26

1	0.0500	2	0.6439	4	0.0133	5	0.0682
6	0.2589	7	0.3431				

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	0	21

TABLE IV-10 (cont'd.)

```
STEP NUMBER          3
VARIABLE ENTERED     2

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1   26

  2    0.6439          3    8.0602          8  54.1304

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1   25

  1    0.3497          4    0.8225          5    1.2859          6    0.4960

  7    0.9414

                 NUMBER OF CASES CLASSIFIED INTO GROUP -
                 Q EQ    X EX
GROUP
  Q EQ             9       0
  X EX             0      21


STEP NUMBER          4
VARIABLE ENTERED     5

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1   25

  2    1.8664          3    4.7101          5    1.2859          8   49.9242

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1   24

  1    0.1152          4    0.0005          6    0.1180          7    0.5118

                 NUMBER OF CASES CLASSIFIED INTO GROUP -
                 Q EQ    X EX
GROUP
  Q EQ             9       0
  X EX             1      20


STEP NUMBER          5
VARIABLE ENTERED     7

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1   24

  2    2.0666          3    5.1251          5    0.8372          7    0.5118

  8   46.8682

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1   23

  1    0.0515          4    0.1484          6    0.0444

                 NUMBER OF CASES CLASSIFIED INTO GROUP -
                 Q EQ    X EX
GROUP
  Q EQ             9       0
  X EX             1      20
```

TABLE IV-10 (cont'd.)

STEP NUMBER 6
VARIABLE ENTERED 4

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 23

2	2.1392	3	4.9303	4	0.1484	5	0.0233
7	0.6415	8	29.5269				

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 22

1	0.1121	6	0.0013

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	1	20

STEP NUMBER 7
VARIABLE ENTERED 1

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 22

1	0.1121	2	1.3623	3	4.7352	4	0.2050
5	0.0668	7	0.6063	8	28.3561		

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 21

6	0.0018

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	Q EQ	X EX
Q EQ	9	0
X EX	1	20

TABLE IV-11
EVENT CLASSIFICATION PROBABILITIES

A classification probability, P, which is less than 0.5 indicates that the event is misclassified with probability 1-P.

Linear Analyses

Earthquakes

		Probability of correct classification				
Event	7 Energy ratios	7 Energy ratios	7 Energy ratios with M_s/m_b set no. 1	2	3	4
Baja Calif.	0.81	0.86	1.00	1.00	1.00	1.00
Baja Calif.	0.74	0.79	1.00	1.00	1.00	0.98
Boxelder Creek	0.33					
Bridgeport	0.97	0.95	1.00	1.00	1.00	1.00
Cache Creek	0.63	0.47	1.00	1.00	1.00	0.99
Cache Creek AS	0.95	0.95	1.00	1.00	1.00	0.99
Colona	0.99	0.98	1.00	1.00	1.00	1.00
Mont.-Wyo. Border	0.99					
Pierre S. Dakota	0.94					
Red Rock River	0.83					
Sierra de Juarez	0.70	0.91	1.00	1.00	1.00	0.99
Teton County	0.64					
W. Maryland	0.91					
W. Vermont	0.96					
28 Jan 62	0.89					
28 Jan 62	0.95					
31 Jan 62	0.92					
24 Apr 62	0.81					
20 Aug 62	0.76	0.78	1.00	1.00	1.00	1.00
16 Nov 62	0.71	0.56	1.00	0.99	1.00	0.59

TABLE IV-11 (cont'd.)

Underground Explosions

Event	7 Energy ratios	Probability of correct classification 7 Energy ratios	7 Energy ratios with M_s/m_b set no. 1	2	3	4
AARDVARK	0.96	0.96	1.00	1.00	1.00	0.99
AGOUTI	0.84	0.89	1.00	1.00	1.00	0.99
ARMADILLO	0.99					
CHINCHILLA II	0.99	0.99	1.00	1.00	1.00	0.99
CIMARRON	0.85					
CODSAW	0.42	0.35	1.00	1.00	1.00	0.99
DANNY BOY	0.56	0.64	1.00	1.00	1.00	0.99
DES MOINES	0.97	0.94	1.00	1.00	1.00	0.99
DORMOUSE II	0.81	0.81	1.00	1.00	1.00	0.99
FISHER	0.81	0.85	1.00	1.00	1.00	0.99
GNOME	0.65					
HARDHAT	0.99	0.95	1.00	1.00	1.00	1.00
HAYMAKER	1.00	1.00	1.00	1.00	1.00	1.00
MAD	0.67	0.85	1.00	1.00	1.00	1.00
MADISON	0.81	0.83	1.00	1.00	1.00	1.00
MARSHMALLOW	0.75	0.79	1.00	1.00	1.00	1.00
MISSISSIPPI	0.89	0.90	1.00	1.00	1.00	0.99
PACKRAT	0.72	0.72	1.00	1.00	1.00	0.99
PAMPAS	0.96					
PASSAIC	0.61	0.62	1.00	1.00	1.00	0.40
PLATTE	0.71					
SACRAMENTO	0.92	0.94	1.00	1.00	1.00	1.00
SMALL BOY	0.71	0.80	1.00	1.00	0.98	0.99
STILLWATER	0.52					
STOAT	0.60	0.15	1.00	1.00	1.00	0.97
WICHITA	0.57	0.67	1.00	1.00	1.00	1.00
YORK	0.56	0.54	1.00	0.99	1.00	0.93

TABLE IV-11 (cont'd.)

Quadratic Analyses

Earthquakes

		Probability of correct classification				
	7 Energy	7 Energy	7 Energy ratios with M_s/m_b set no.			
Event	ratios	ratios	1	2	3	4
Baja Calif.	0.99	1.00	1.00	1.00	1.00	1.00
Baja Calif.	1.00	1.00	1.00	1.00	1.00	1.00
Boxelder Creek	0.05					
Bridgeport	0.99	1.00	1.00	1.00	1.00	1.00
Cache Creek	0.46	0.74	1.00	1.00	1.00	1.00
Cache Creek AS	0.99	0.99	1.00	1.00	1.00	1.00
Colona	0.99					
Mont.-Wyo. Border	0.99					
Pierre S. Dakota	0.99					
Red Rock River	0.94					
Sierra de Juarez	1.00	1.00	1.00	1.00	1.00	1.00
Teton County	0.83					
W. Maryland	1.00					
W. Vermont	0.99					
28 Jan 62	0.99					
28 Jan 62	1.00					
31 Jan 62	0.98					
25 Apr 62	0.91					
20 Aug 62	0.99	0.99	1.00	1.00	1.00	1.00
16 Nov 62	0.65	0.84	1.00	1.00	1.00	1.00

TABLE IV-11 (cont'd.)

Underground Explosions

Event	Probability of correct classification: 7 Energy ratios	7 Energy ratios	7 Energy ratios with M_s/m_b set no. 1	2	3	4
AARDVARK	1.00	1.00	1.00	1.00	1.00	1.00
AGOUTI	0.97	1.00	1.00	1.00	1.00	1.00
ARMADILLO	1.00					
CHINCHILLA II	1.00	1.00	1.00	1.00	1.00	1.00
CIMARRON	0.98					
CODSAW	0.99	0.99	1.00	1.00	1.00	1.00
DANNY BOY	0.69	0.98	1.00	1.00	1.00	1.00
DES MOINES	1.00	1.00	1.00	1.00	1.00	1.00
DORMOUSE II	0.99	1.00	1.00	1.00	1.00	1.00
FISHER	0.88	0.99	1.00	1.00	0.99	0.99
GNOME	1.00					
HARDHAT	1.00	1.00	1.00	1.00	1.00	1.00
HAYMAKER	1.00	1.00	1.00	1.00	1.00	1.00
MAD	0.90	0.95	1.00	1.00	1.00	1.00
MADISON	0.99	1.00	1.00	1.00	1.00	1.00
MARSHMALLOW	0.94	1.00	1.00	1.00	1.00	1.00
MISSISSIPPI	0.99	1.00	1.00	1.00	1.00	1.00
PACKRAT	0.99	1.00	1.00	1.00	1.00	1.00
PAMPAS	1.00					
PASSAIC	0.79	1.00	1.00	1.00	1.00	1.00
PLATTE	0.89					
SACRAMENTO	0.99	1.00	1.00	1.00	1.00	1.00
SMALL BOY	0.99	1.00	1.00	1.00	1.00	1.00
STILLWATER	0.97					
STOAT	1.00	1.00	1.00	1.00	1.00	1.00
WICHITA	0.71	0.94	1.00	1.00	1.00	1.00
YORK	0.66	1.00	1.00	1.00	1.00	1.00

TABLE V-1
SUMMARY OF F-STATISTICS FOR EACH VARIABLE

The F-statistic may be used to test the hypothesis that the two group means are unequal.

The threshold F values required to reject the hypothesis that the two group means are equal at the 99 per cent confidence level are:

$F \geq 7.31$ (on 1 and 45 degrees of freedom)
$F \geq 7.64$ (on 1 and 28 degrees of freedom)

Variable Number	Variable Description		F (47 Events) (1,45) D.F.	F (30 Events) (1,28) D.F.
1	P_1/S_1		15.52	6.41
2	P_2/S_2		6.91	2.57
3	Pg/B		16.54	6.77
4	Pg/Lg_1		20.55	9.25
5	$Pg/(Lg_1+Rg_1)$		12.35	6.25
6	Lg_1/Rg_1		0.32	0.08
7	Lg_2/Rg_2		0.01	1.23
8*	M_s/m_b (M_s-m_b)	set 1		214.56 (225.09)
8	" "	set 2		115.50 (115.45)
8	" "	set 3		74.78 (80.43)
8	" "	set 4		43.61 (51.83)

* For comparison purposes, the F-statistics for M_s-m_b are shown in parentheses.